THE PATH TO THE CONCEPTION OF THE JUNCTION TRANSISTOR

William Shockley tells how he and his associates conceived and developed the junction transistor for which he, J. Bardeen and W. Brattain, were awarded a Nobel Prize in 1956. He emphasizes the importance of creative failure and of the will to think. He "hopes that the presentation of details of his limitations in making this important invention may help readers to accept their own limitations and, thereby, to become more persistent and, hence, creative."

Reprinted by permission from the copyrighted IEEE Transactions on Electron Devices, July 1976, Vol. ED-23 No. 7, which contains 19 other case histories of inventions, told by the inventors.

IEEE TRANSACTIONS ON

ELECTRON DEVICES

JULY 1976 VOLUME ED-23 NUMBER 7

A PUBLICATION OF THE IEEE ELECTRON DEVICES SOCIETY

SPECIAL ISSUE: HISTORICAL NOTES ON IMPORTANT TUBES AND SEMICONDUCTOR DEVICES

The Path to the Conception of the Junction Transistor

WILLIAM SHOCKLEY, FELLOW, IEEE

***Abstract*—The failure in 1945 of experiments proposed by Shockley, on what today would be called thin-film field-effect transistors, was a creative failure that stimulated Bardeen in early 1946 to propose that a surface-state shield blocked the field from the semiconductor's interior. Bell Laboratories' "transistor group to be" for the next eighteen months focused, not on practical, hut on scientific aspects of the failure. Focus on the practical resumed (with a step-function increase, lasting several months, in "the will to think" about new concepts of semiconductor amplifiers, as measured by the rate of filling of laboratory notehook pages by Bardeen, Brattain, and Shockley) on 17 November 1947, when in his surface-state research, Brattain penetrated the shield by applying the field through an electrolyte. Within six days, patentable field-effect transistor inventions were conceived. Although useless as devices, these inventions were creative failures used hy Bardeen and Brattain to discover the point-contact transistor three weeks later. Five weeks after this discovery, Shockley conceived the junction transistor while designing "imref" experiments on the point-contact transistor's inversion layer so that in 1951, the point-contact transistor in its turn became a creative failure when replaced hy the junction transistor whose conception it had aided. But the path of thought to the conception of the junction transistor and the subsequent path to its practical realization are proven to be highly indirect by historical research on lahoratory notebook entries. Specifically, Shockley's conception of the junction transistor was delayed by at least four months because he missed opportunities, obvious by hindsight, to recognize the possibility of minority carrier injection. The author hopes that the presentation of details of his limitations in making this important invention may help readers to accept their own limitations and, thereby, to become more persistent and, hence, creative.**

In the foyer of the main entrance to Bell Laboratories, along with a bust of Alexander Graham Bell, there appears the following statement credited to him:

> Leave the beaten track occasionally and dive into the woods. You will be certain to find something that you have never seen before.

I. Introduction: "Creative-Failure Methodology"

The EDITORS of this Bicentennial Issue urged me in this historical note to focus on the junction transistor. At first, I thought that my chief purpose would be to analyze and to present the sequence of thoughts, the interactions among individuals, and the motivations that led to the conception of the junction transistor. But as I wrote—and rewrote—my manuscript, I found that what would convince me that my efforts on this article had been worthwhile would be accidentally to overhear a conversation when someone said: "You know, I might have quit on my research that finally paid off, if I hadn't read Shockley's article on how slow he was and how he missed the junction transistor's key concepts so many times. If he was that dumb, I figured I should stick with it and not give up."

The editors' request for emphasis on the junction transistor is in keeping with the announced objective of this Bicentennial Issue to present "Historical Notes on Important Tubes and Semiconductor Devices," if "important" is interpreted as "technologically important." Technologically important does not apply to the emphasis of many histories about the invention of the transistor which stop on Christmas Eve of 1947. That was the date of "reduction to practice" of the point-contact transistor in Walter Brattain's famous notebook entry about the previous day's demonstration of the point-contact transistor for the leaders of the Research Department at Bell Laboratories (then called Bell Telephone Laboratories).

My conception of the junction transistor took place about one month after the discovery of the point-contact transistor. The point-contact transistor was the device under development at Bell Laboratories for about three years. In early 1951, it was displaced—after the obvious superiority of the first microwatt junction transistors was clearly demonstrated. Since then, junction transistors, fabricated by a wide diversity of processes and incorporated into integrated circuits, have continued to play a major role in solid-state electronics and, thus, are the first of the technologically important devices of the solid-state era.

The junction transistor focus of this article extends the theme of "creative-failure methodology" of my three prior papers on transistor history. All three papers emphasize how the failures of attempts to make field-effect transistors became "creative failures" by creating the program that discovered the point-contact transistor. The titles of these papers indicate this emphasis: "Transistor History, Applied Research and Science Teaching," in 1963 [1], and "The Invention of the Transistor: An Example of Creative-Failure Methodology," for both 1973 [2] and 1974 [3]. The new emphasis of this article is that in terms of technological significance, the point-contact transistor became a creative failure by setting up challenging scientific problems. My response to these challenges was what finally led me to the conception of the junction transistor.

Among many other historical reviews, the one which is most relevant to my discussion here is that of John Bardeen and Walter Brattain in their 1949 *Physical Review* article [4]. Our three Nobel lectures [5]–[7] are also important sources, and, of course, so are the three letters to the editor of the *Physical Review* [8]–[10] published at the time of the public announcement of the transistor. The chief historical references to the junction transistor are my 1949 *Bell System Technical Journal* article [11], my 1950

Manuscript received March 22, 1976.
The author is at 797 Esplanada Way, Stanford, CA 94305.

book *Holes and Electrons in Semiconductors* [12], the Shockley, Sparks, and Teal 1951 *Physical Review* article [13], and my 1951 junction transistor patent [14]. Walter Brattain's recent article "Discovery of the Transistor Effect" [15] discusses conferences among members of the semiconductor group leading to his and my collaborative first measurement of the density of Bardeen's surface states and his own suggestion of the surface photovoltaic effect that set the scene for the key observation, discussed in Section VI, that led within one month to the birth of the point-contact transistor. However, I question (see Section VIII) the historical accuracy in [15] of some features of Brattain's "discovery recollections."

Because I refer to positrons while discussing holes in lectures, I was stimulated to do historical research by Brattain's speculation in [15]: "Hindsight makes me wonder whether the analogy between the creation of a hole–electron pair in a semiconductor and an electron–positron pair in free space might have influenced Wilson." I found this date sequence: In 1930, Dirac's book interpreted a "hole" in his negative energy states, not as a positron, but as a proton. In lectures, I state that the high mobility of holes can't be made obvious: What Dirac missed is by definition nonobvious. In 1931, Wilson's theory of holes and electrons in semiconductors was published. In 1932, the positron was discovered experimentally by Anderson. Anderson's 1961 lecture on related history credits Blackett and Occialini as first suggesting in 1933 the pair-production hypothesis.

Rather than citing the reference for every publication mentioned below, I shall in most cases simply refer to one or more of the above references where citations may be found. I shall apply this specifically to a number of the publications that were precursors to the junction transistor references [11]–[14]. This procedure does not adequately recognize the contributions of the many workers who laid the foundations upon which the transistor program was based. For historical perspective that does include appropriate acknowledgments, the reader is referred in particular to Bardeen's Nobel lecture [5] and also to [4], [6], [12], [15], etc.

For this article, as for my 1974 article [3], I did research, reminiscent of my World War II operations research, on available documentary records. This research enabled me to obtain information and with it to reconstruct portions of the history with an accuracy that would have been impossible on the basis of my own memories and my personal records. I discovered striking quantitative data that gave perspective on the stimulations and the motivations of the transistor group during the period of peak creativity at the end of 1947 and early 1948. I shall use "transistor group" to signify those who became members of the small confidential group who knew about the transistor's existence prior to the public announcement on 30 June 1948. Actually, the word "transistor" was not used until several months after the device was born. Walter Brattain describes how John R. Pierce invented the name during a conference in Brattain's office [15].

Based on my operations research experience with these documents, I conclude that laboratory notebooks are reliable in establishing facts—some of them now as much as 36 years old—about the history of the transistor program. There are obvious reasons for this. The date in a notebook of the disclosure by one inventor of his invention to a competent witness, who reads and understands it, may be decisive in the patent being issued to him.[1]

At Bell Laboratories, every notebook is numbered and a scrupulous record shows to whom it was assigned and when. These procedures ensured, with few exceptions, the availability of the notebook records that I needed for my research on this historical note. I have used some of the entries that I found important as figures here—including, of course, the most famous of all, Walter Brattain's Christmas Eve report of 1947.

That notebook entry underlines the significance of one source of the motivations in the transistor group. It is, of course, evident that motivation may be based upon a wide variety of factors. Not the least of these, as reflected in Brattain's notebook entry, is based on attitudes about patents. Even without a financial reward, the prestige factor of an authentic demonstration of an inventive contribution can motivate the obtaining of a patent.[1]

This historical note has "creative-failure methodology" as a central theme. I discuss its role in the interactions among the individuals of the transistor group, including the internal competition in which I myself was involved, and its relevance to human limitations—particularly my own slowness in recognizing the key concepts of the junction transistor. I shall give examples of how an acceptance of one's limitations stimulates creativity by showing that failures need not be simply accepted as causes for low self-esteem but can instead be used as stepping stones on the path of progress. To convey the spirit of creative-failure methodology is one of my principal purposes in this article. This purpose is, so to speak, a double-negative contribution to creativity. It is intended to negate a negative attitude: An understanding of creative-failure methodology may not per se inspire creativity but appreciation of its spirit may remove barriers to progress after one leaves the beaten track and dives into unfamiliar woods.

II. Three Key Concepts of the Junction Transistor

Three key concepts that I shall specially emphasize as essential to the junction transistor are:

1) minority carrier injection into the base layer which increases exponentially with forward emitter bias,

2) application of reverse voltage at the collector junction,

[1] Our nation's founders motivated invention by devising the constitutional power of Congress "to promote the progress of science and the useful arts, by securing for limited times to authors and inventors and exclusive right to their respective writings and discoveries." Unsound court decisions in patent cases may be degrading the value of patents and thereby frustrating the incentive of patents summed up in the acronym IBLER (incentives based on limited exclusive rights). See W. Shockley's testimony in "Technology and Economic Growth", Hearings before the Subcommittee on Economic Growth of the Joint Economics Committee, July 15 and 16, 1975 for sale by the Superintendent of Documents, U.S. Gov. Printing Office, Washington, DC 20402, price $2.40.

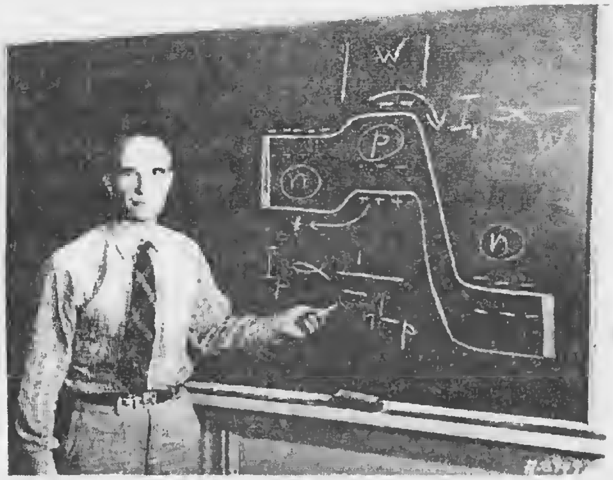


Fig. 1 The author presenting the theory of the junction transistor probably during the "persistor" period of 1949 or early 1950.

3) favorable geometry and doping levels so as to obtain good emitter to collector efficiency.

These three key concepts of the junction transistor, except for the exponential voltage dependence, are compactly introduced in historical perspective by Fig. 1.[2]

In the diagram for the n-p-n transistor of Fig. 1, the emitter junction is forward biased as indicated by the lower Fermi level (or "imref" as discussed in Section III) in the p-type base layer. Minority electrons are shown diffusing through the base layer which is substantially at uniform potential. The collector has a large reverse bias—a central concept in obtaining voltage gain through a high output impedance. The unwanted hole current from base into emitter is inversely proportional both to the emitter conductivity (shown on the blackboard at the tip of the piece of chalk) and to the diffusion length of holes in the emitter. The useful current to the collector and, hence, the emitter efficiency are both inversely proportional to the base-layer thickness.

I discovered the photograph of me in Fig. 1 in the files of Bell Laboratories' Publications Department during research that I did in 1972 for an invited paper [2] for the Second European Solid State Device Research Conference held at the University of Lancaster, England, for the 25th anniversary of the transistor. I infer the date of 1949 or 1950 for Fig. 1 on the basis of my seriously determined expression. My reasoning is as follows:

The date of conception of the junction transistor was 23 January 1948. I wrote a detailed theoretical analysis in a patent application filed on 26 June 1948 [14] and published an extended one in mid-1949 [11]. In April 1949, some very inadequate "existence-proof" amplifying devices were made. These did not demonstrate the technological potential of the device. At about that same time, I remember that Bert Moore, an electrical engineer in our group, suggested that the name "persistor" would be appropriate for the junction transistor, if it ever came to be useful, because persistence was certainly what was needed to produce it. A satisfactory realization of the junction transistor was not achieved until the Spring of 1950 and real excitement about it did not develop until early 1951. I believe that my facial expression in Fig. 1 expresses my attitude towards the disappointments that characterized the "persistor" period of late 1949 and early 1950.

III. "The Will to Think"

Before I reviewed my notebook entries for the purpose of writing this historical note, I had the impression that all of the concepts needed to create the junction transistor combination shown in Fig. 1 had been *well formulated* many months earlier. The record, that I discuss below, shows that I was aware of the general considerations needed to develop the theory of the junction transistor. But I did not formulate several key concepts until after the junction transistor was invented. Before that, my analysis went only far enough to show how experiment and theory on p-n junctions might interact so as to add scientific knowledge about the basic physical phenomena. Also I found that I had speculated about several uses of p-n junctions in practical devices. But none of these speculations were analytically developed.

What motivated *the will to think* through the logical, sometimes mathematical, relationships needed for the invention of the junction transistor was finally provided chiefly by two factors; first, my own motivation to play a more significant personal, rather than managerial, role in what was obviously a development of enormous potential importance, and second, the challenge to resolve some of the puzzles about the operation of the point-contact transistor—and chance also played a role.

"The will to think" phrase expressively describes the peak of inventive activity that was initiated by a specific event that occurred on 17 November 1947. During the following month, the point-contact transistor was invented and demonstrated as a working amplifier. During that same month—labeled the "magic month" in my 1973 paper—three other closely-related inventions were made. Somewhat later, on 23 January 1948, I invented the junction transistor. These five inventions were the subject of the first transistor device patent applications filed by Bell Laboratories. All five were filed before the first public announcement of the transistor on 30 June 1948. (See [3] for details on these patents.)

I first heard the phrase "the will to think" from Professor Enrico Fermi in 1940. Jim Fisk and I went to visit him to discuss his research on atomic energy. Bell Laboratories was also exploring the possible military significance of nuclear fission in response to a request from the National Academy of Sciences.

Fermi was designing experiments to study the slowing down of neutrons in graphite. He had confidence that such experiments would be carried out because financial support by the U.S. Government had finally been assured to the project. He said that this assurance gave him "the will

[2] In discussing Fig. 1, I shall assume that the reader is familiar with the basic concepts and terminology of semiconductor electronics including energy bands, energy gap, electrons, holes, donors, acceptors, minority carrier lifetime, emitter, base, collector, etc. See, for example, [1], [2], [12].

to think." In these four words he distilled the essence of a very significant insight: A competent thinker will be reluctant to commit himself to the effort that tedious and precise thinking demands—he will lack "the will to think"—unless he has the conviction that something worthwhile will be done with the results of his efforts—and, of course, there is always also the risk that his hard thinking may not produce any creative ideas. This was the reasoning, that I remember Fermi told me, why assurance of financial support gave him the will to try to think through the optimum design for his experiments.

"Try simplest cases" is a phrase that contributes to the will to think. I formulated it many years after my visit with Fermi while doing research on science teaching. I have defined it as a "search thinking tool" applicable at the high school level [16], [17]. It is based on my own research and teaching experiences including those in the transistor program. A meaningful simplest case stimulates the will to think by reducing the threat of being forced to accomplish repugnant and tedious tasks.

Our visit to Fermi was motivated by a "simplest case" invention, made by Jim Fisk and me, of a certain structure. That structure was a mathematical model of a nuclear power plant that we had conceived during our research on atomic energy. Fermi and his colleague Leo Szilard had, I believe, independently made the same invention—probably even earlier. The calculations that supported the merit of the invention involved neutrons slowing down and diffusing in graphite or other moderator. The use of the theory of diffusion for this purpose was like what I did several years later in analyzing the junction transistor.

I cannot now recall whether my 1940 experience with neutrons diffusing in slabs of graphite influenced my 1948 thinking of minority carriers diffusing through a transistor's base layer. The mathematics is closely similar for the two cases. But I have no doubt that my ability to contribute to both programs was strengthened by my using the try-simplest-cases approach.

The junction transistor was a simplest case. It avoided many complexities because of its essential simplicities: Its analysis was one-dimensional and linear! So were the calculations for the structure that I had in mind when visiting Fermi. That structure utilized the basic invention of "segregating" or "lumping" the fissionable material so as to lessen the need for enriching the active uranium isotope to obtain a chain reaction. Fisk and I first did our calculations for the simplest geometrical case of parallel layers, the "*L* Case." Only after we had obtained encouraging results for this simplest case did we have the will to think through the more complex cases of parallel cylinders or of a lattice of spheres.

Fisk and I put our simplest-case mental picture into the cartoon form shown in Fig. 2 for the case of hydrogen, rather than carbon, as moderator. For many years this cartoon was classified and is published here for the first time. These same security considerations were probably what kept Fermi from exchanging any ideas with us about segregation during our conference. The report that Fisk

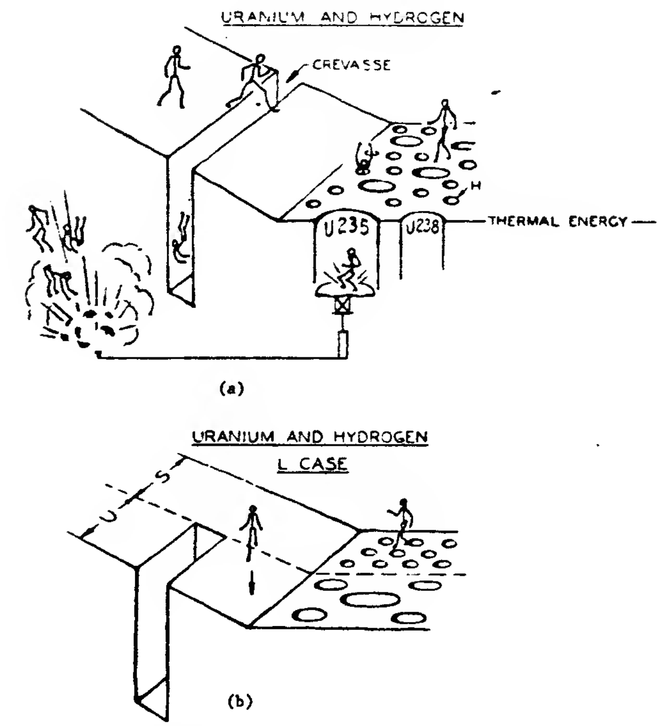


Fig. 2. The theory of segregation as visualized by J. B. Fisk and W. Shockley in 1940 for the "simplest case" of parallel slabs in a hydrogen-moderated nuclear reactor. A neutron of thermal energy captured by isotope U235 will cause several high energy neutrons to be generated during fission. These neutrons lose energy by collisions with protons and are threatened by useless capture by U238 at a resonance (crevasse) energy of a few volts. If they do reach thermal energy, they can diffuse among the hydrogen atoms, which have small capture cross sections, and be captured by U235 or U238. A chain reaction requires sufficient probability of U235 capture. This probability can be increased by "enriching" the abundance of U235 and also by decreasing resonance capture by segregating the uranium into *U* layers so that some neutrons avoid the crevasse by slowing down in *S* layers and later by diffusing in the *S* layers to reach the uranium. (a) Shows the homogeneous case and (b) the layer or "*L* case."

and I prepared contributed, I have been told, to some aspects of the atomic energy program, especially in England. Bell Laboratories filed for patents on our work but, not being in the power business, undertook no other activity. (The junction transistor story is not concerned with the atomic energy patent applications.)

I have published several research papers based on simplest cases, not related to diffusion. Some of my more recent ones have the "try simplest cases" phrase in their titles. My earliest such publication was the "empty lattice" test of John Slater's cellular method of calculating wave functions in crystals. This research was a by-product of my Ph.D. thesis. I find most satisfaction, perhaps, in the "simplest-case" research result that I obtained in the 1970's as a deliberate test of my science teaching methodology. It made use, in part, of parallel slabs—faintly reminiscent of the "*L* Case" and of the junction transistor. The result was my resolution of a problem that had puzzled Nobel Laureates Pauli, Tamm, and von Laue, and perhaps even Einstein, by establishing very directly the correct formula for electromagnetic momentum in matter [18].

Another thinking tool that played a key role in my junction transistor analysis is the Fermi level, or its

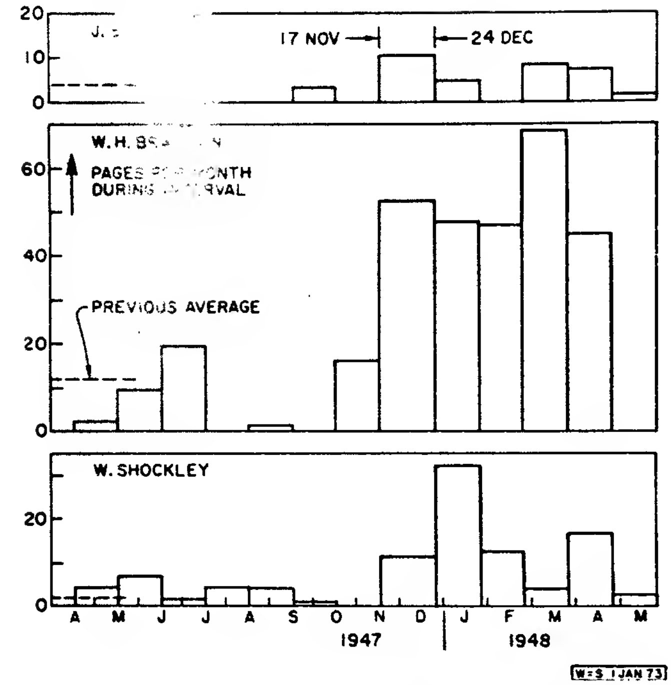


Fig. 3. The quantitative impact of the observation of 17 November 1947 on "the will to think" as shown by the number of notebook pages used per month. (Areas represent total pages used and ordinates represent rates. Each bar is one calendar month except for the one starting on 17 November 1947 which is lengthened to Christmas Eve 1947 to include the famous notebook entry of Walter H. Brattain reproduced in Fig. 17. After 25 April 1948, Brattain shared notebook use with an increased number of technical assistants; since the data then were not comparable to earlier data, they have been omitted.)

pseudo-equilibrium version, the quasi-Fermi level. Considerations directly related to differences in quasi-Fermi levels were what led me to the mental pictures that completed the concept of the junction transistor. Quasi-Fermi levels significantly simplified junction transistor theory and lent conviction to the validity of the concept and, thereby, contributed to the will to think of ways to make junction transistors become a reality.

I once discussed the great merit of quasi-Fermi levels with Fermi himself and asked him for a substitute that would reduce their cumbersome six-syllable length. He suggested "imref." I used "imref" in the microwatt junction transistor paper [13]. A footnote suggested that, if puzzled, as I had been, the reader should spell it backwards, as Fermi had told me to do. I shall use "imref" in preference to quasi-Fermi level in the remainder of this article.

Fig. 3 displays the quantitative impact of the increase in the will to think that led to the inventions of the point-contact and junction transistors. It is based on a review of the laboratory notebooks assigned to John Bardeen, to Walter Brattain, and to me. The "step function" increase in activity shown in Fig. 2 is expressed in the number of pages filled per month for each of the three of us. During the first week following the key date of 17 November 1947, Brattain used 15 pages and Bardeen used 10 pages compared, respectively, to 8 and 0 the week before.

These rates of filling pages for the week starting on 17 November 1947 compared to long-term averages, were more than 4 times greater for Brattain and more than 10 times for Bardeen. That the highlevel activity of the 17 to 24 November week was the initiation of a trend is clear in Fig. 3; the persistence of that high level that started on 17 November continued through the whole first quarter of 1948. The narrative of events, presented below, shows in detail why this occurred. My own increased upward trend, lasting for many months, started 3 weeks later on 8 December, and in the following week, I used 5 pages—about 10 times my long-term weekly average.

What was the event on 17 November 1947 that initiated this dramatic stepped-up tempo? Why did it so stimulate "the will to think" about items worth entering as permanent records in our laboratory notebooks?

My answers to these questions are presented in the narrative of events that constitutes the major portion of this historical note. For ease of reference, I have made a table of relevant dates. Those dates that I believe most significantly influenced my thinking on the path to the conception of the junction transistor appear in bold face type. Many of the others relate to cases when I made a notebook entry that might have led to the invention of the junction transistor—in effect, opportunity knocked but was not heeded.

JB, WHB, and WS in Table I indicate John Bardeen, Walter H. Brattain, and William Shockley in whose assigned laboratory notebook the NBE (notebook entry) appears. The time period from late 1939 to early 1951 is divided into five periods as indicated by headings of Sections IV through VIII, where the individual items are discussed.

IV. History Before World War II

Research for my Ph.D. at the Massachusetts Institute of Technology later contributed to my motivation toward solid-state electronics at Bell Laboratories. I was one of Professor Slater's students when his chief interest emphasized wave functions in crystals. My thesis involved sodium chloride. About fifteen years later, I saw that a feature of the band edges for my thesis calculations must represent a general situation and published a brief note applying it to electron and hole masses in connection with transistor-related semiconductor theory.

I received my Ph.D. in 1936 and went to Bell Laboratories the same year. I was assigned to report to Dr. C. J. Davisson. Dr. Davisson and his colleague, Dr. Lester Germer, at Bell Laboratories, had first observed electron diffraction. After being formally assigned to Dr. Davisson, I was put on loan to the Vacuum Tube Development Department and worked on electron multiplier tubes and other problems on vacuum-tube theory. These experiences oriented me to practical electronic problems.

A key motivation that stimulated my will to think about transistors came from Dr. Mervin Kelly, who was then the Director of Research at Bell Laboratories, a position he held before becoming President some years later. Dr. Kelly visited me for the purpose of emphasizing his objective of introducing electronic switching in the telephone system. He said that he looked forward to the time when metal

TABLE I

A List of Dates with Brief Descriptions of Important Events

Especially important dates from the author's viewpoint on the path to the conception of the junction transistor appear in bold face type. The numbers and titles of the five periods match those of the sections where they are discussed with the dates indicated by bold face or *italics* as in this table.

	IV. PRE-WORLD WAR II
29 Dec 39	WS NBE: disclosure of what would now be called "Schottky-gate field-effect transistor."
29 Feb 40	WS NBE: improved version of 29 Dec 39.
7 Jun 40	WS NBE: disclosure of "*L* Case" uranium segregation for reactor without enrichment of U235 (see Fig. 2).
	V. FROM 1945 UP TO 17 NOVEMBER 1947
16 Apr 45	WS NBE: field-effect idea using a p-n junction.
23 Jun 45	WS NBE: quantitative estimate of the degree of failure of thin-layer field-effect experiments.
19 Mar 46	JB NBE: proposal of surface states to explain failure of 23 Jun 45.
12 Mar 47	WS NBE: very preliminary theory of p-n junction resistance.
4 Apr 47	WS NBE: lightning arrestor using alternating p- and n- type layers.
24 Apr 47	WS NBE: diffusion theory of minority current to a reverse-biased p-n junction.
19 May 47	WS NBE: theory of reverse currents in p-n junctions including Zener effect.
16 Sep 47	WS NBE: n-p-n structure for a high-frequency thermistor involving electron flow over the potential maximum of the p-layer.
	VI. THE "WILL TO THINK" AND THE BIRTH OF THE POINT-CONTACT TRANSISTOR
17 Nov 47	WHB NBE: overcoming surface states using Gibney's suggestion of changing bias of the electrolyte.
4 Dec 47	WHB NBE: success of several device ideas including WS proposal to modulate p-n junction resistance with voltage applied to a drop of electrolyte over the junction.
8 Dec 47	WS NBE: concepts of a junction, field-effect transistor and of voltage gain by using reverse bias on a p-n junction, the latter being key concept (2) of Section II.
8 Dec 47	WHB NBE: achievement of voltage and power gain by using reverse voltages on "high-back voltage" germanium.
16 Dec 47	WHB NBE: first record of voltage and power gain in a point-contact transistor.
	VII. THE CONCEPTION OF THE JUNCTION TRANSISTOR
31 Dec 47	WS NBE: a p-n-p structure almost—but not quite—involving minority carrier injection into a base layer.
23 Jan 48	WS NBE: *the conception of the junction transistor*, including concepts (1) and (3) of Section II.
	VIII. FROM CONCEPTION TO REALIZATION OF THE JUNCTION TRANSISTOR
18 Feb 48	J. N. Shive reported at a conference his observations on the double-surface transistor.
7 Jun 48	J. R. Haynes NBE: first version of Haynes-Shockley experiment showing that a positive point injects holes into n-type germanium that drift in an electric field.
26 Jun 48	WS: patent application [14] filed for junction transistor.
1 Dec 48	E. J. Ryder and WS published as letter (received 1 Dec 48) by the *Physical Review* reporting hole injection lowers bulk resistance of n-type germanium. (The experiment was done several months earlier.)
7 Apr 49	"Existence proof" for the junction transistor. Made by dropping molten p-type germanium onto hot n-type and sawing the resulting junction to make two strips of p-type tying on an n-type plane. Preparation by R. M. Mikulyak under the direction of M. Sparks [3].
49 and 50	Good p-n junction prepared by M. Sparks and G. K. Teal by changing the doping of the melt in a germanium crystal grower by "pill dropping" and subsequent double doping to make n-p-n structures. The development and publication of junction transistor theory [11], [12] added momentum to the program.
Early 51	Realization and demonstration of the first micro-watt, grown-junction transistors.

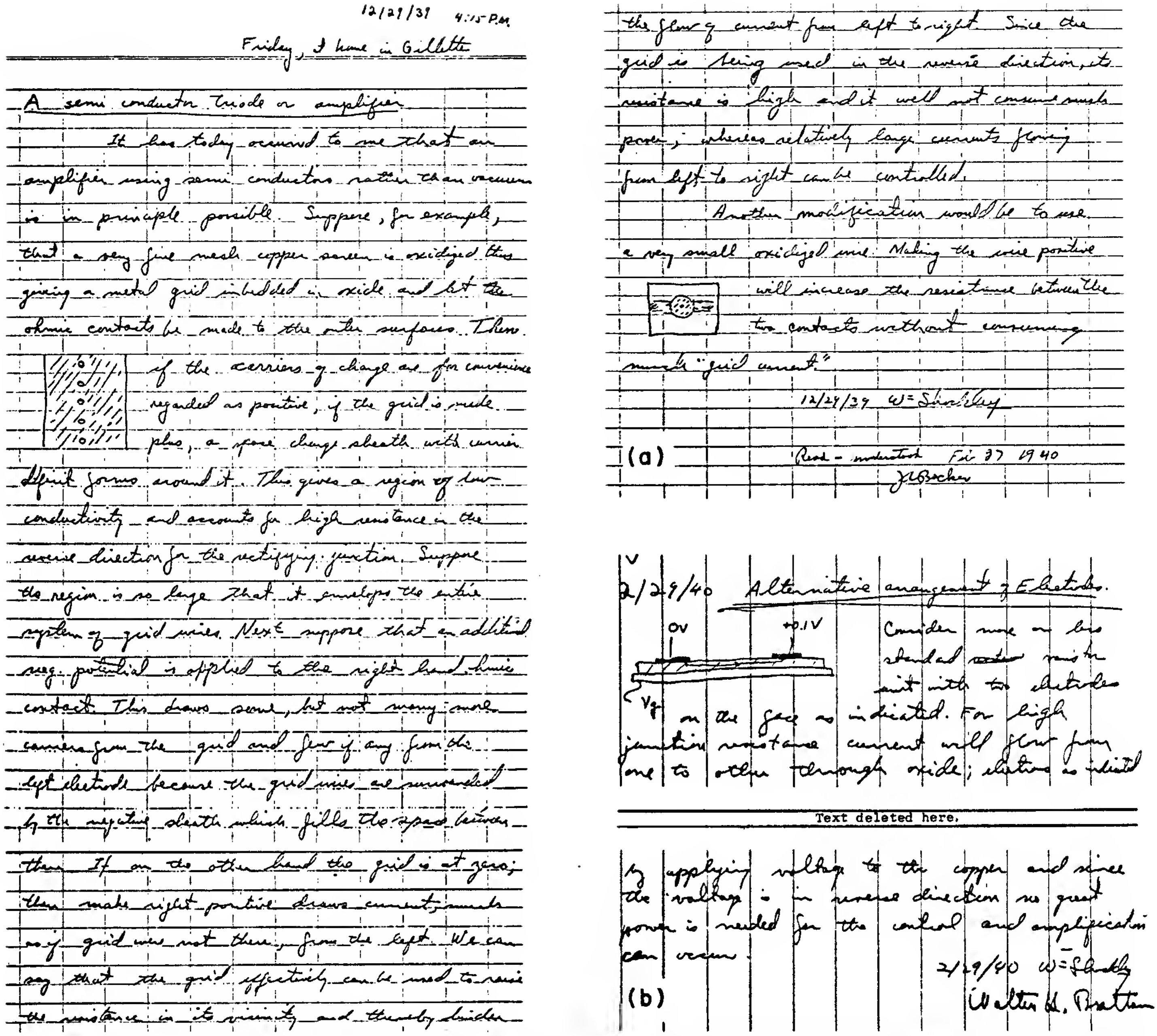
contacts, which were used in telephone exchanges to make connections when numbers that are dialed, would be replaced by electronic devices. His interest in this goal was very great. He stressed its importance to me so vividly that it made an indelible impression.

My interests in the vacuum-tube field were not as great as my interests in some aspects of solid-state research going on at Bell Laboratories. The atmosphere at Bell Laboratories was such that it was possible for me to change the emphasis of my work and I was permitted to concentrate my efforts in the solid-state field. I worked on the theory of alloys and, in particular, on the theory of order and disorder transformations.

It occurred to me some time around 1938 or 1939 that there would be a possibility of achieving the objective of electronic switching that Dr. Kelly had in mind, using phenomena in solid-state physics rather than the vacuum-tube techniques that he had considered. One possibility that occurred to me was a solid-state amplifier using carbon contacts or some other type of contact subject to pressure that was controlled by an input signal applied to a quartz crystal or some other piezoelectric crystal. The output power would be obtained through the change in resistance of the microphonic contact. Although I did not know it at the time, this was an old idea. Mr. Alan Holden and I made some attempts to make an amplifier this way, but concluded that this approach held very little promise indeed.

At about the same time Dr. Brattain and his supervisor, Dr. J. A. Becker, involved me in their research on copper-oxide rectifiers. This stimulated me to study the theory by Schottky of rectification by metal–semiconductor contacts, a theory now made familiar in the phrases "Schottky barrier" and "Schottky gate." While considering Schottky's theory and having ideas about amplification in the back of my mind, I recognized that possibilities of amplification were inherent in Schottky's depletion layer—the space-charge layer that spreads more deeply into the semiconductor as the reverse potential on the rectifier is increased. I saw that this spreading could be used as a kind of valve action so as to control conductivity in the semiconductor at a substantial distance from the contact.

29 December 1939 is the date of my first notebook entry, reprinted in Fig. 4(a), that proposed a semiconductor amplifier. I regard it as the first milestone on my path to the concept of the junction transistor. I let two



12/29/39 4:15 P.M.
Friday, I home in Gillette

A semi conductor triode or amplifier

It has today occurred to me that an amplifier using semi conductors rather than vacuum is in principle possible. Suppose, for example, that a very fine mesh copper screen is oxidized thus giving a metal grid imbedded in oxide and let the ohmic contacts be made to the other surfaces. Then if the carriers of charge are for convenience regarded as positive, if the grid is made plus, a space charge sheath with carrier deficit forms around it. This gives a region of low conductivity and accounts for high resistance in the reverse direction for the rectifying junction. Suppose the region is so large that it envelops the entire system of grid wires. Next suppose that an additional neg. potential is applied to the right hand ohmic contact. This draws some, but not many more carriers from the grid and few if any from the left electrode because the grid wires are surrounded by the negative sheath which fills the space between them. If on the other hand the grid is at zero; then make right positive draws current, much as if grid were not there, from the left. We can say that the grid effectively can be used to raise the resistance in its vicinity and thereby hinder the flow of current from left to right. Since the grid is being used in the reverse direction, its resistance is high and it will not consume much power; whereas relatively large currents flowing from left to right can be controlled.

Another modification would be to use a very small oxidized wire. Making the wire positive will increase the resistance between the two contacts without consuming much "grid current."

12/29/39 W Shockley

(a) Read - understood Feb 27 1940 J A Becker

2/29/40 Alternative arrangement of Electrodes.

OV +0.1V

Vg

Consider now a standard oxide rectifier with two electrodes on the face as indicated. For high junction resistance current will flow from one to other through oxide; electrons as indicated

Text deleted here.

by applying voltage to the copper and since the voltage is in reverse direction no great power is needed for the control and amplification can occur.

2/29/40 W Shockley
Walter H. Brattain

(b)

Fig. 4. The Shockley notebook disclosures of a Schottky-gate field-effect transistor. (a) The first disclosure of 29 December 1939 discussing the modulation of conductivity by the spreading of Schottky's depletion layer away from a metal–semiconductor junction. (b) An improved structure dated two months later on 29 February 1940.

months slip by at that time before having my disclosure witnessed, as may be seen in Fig. 4(a) by the "read–understood" endorsement of 27 February 1940 by J. A. Becker. Two days after that, I entered the improved structure, shown in Fig. 4(b). (For the complete notebook entry of 4(b), see [3].)

The invention of Fig. 4(b) was theoretically sound. It describes a device of the type now known as a Schottky-gate field-effect transistor. The basic concept of the two disclosures of Fig. 4 is not, as might be supposed from the grid structure shown as part of Fig. 4(a), simply the analog of a vacuum tube with a metal grid inserted in the Schottky depletion layer. But a grid in the depletion layer seems to be what Walter Brattain remembered more than thirty years later in his reminiscences of our collaboration in 1940 as published in [19] for the transistor's 25th anniversary and again later [15]:

> I vividly recall Becker's and my recognition of the close analogy between the copper-oxide rectifier and the vacuum-tube diode, and our calculations of the size of the grid that one might put into the space-charge layer of the rectifier to make a triode! It is an understatement to say that the results did not look promising. So I was somewhat amused when a year or so later, William Shockley came to me with an idea for making an amplifier out of copper-oxide. As I remember, I nevertheless told him that the possibility of achieving this was so important that I would try to make the copper-oxide device he had in mind as near as possible the way he wanted it. It didn't work.

From this I conclude that Brattain's signature on the disclosure of Fig. 4(b) must have been that of an amused

witness rather than that of a serious coinventor: he focused exclusively on a grid in the space-charge layer in his reminiscences. My notebook entries of early 1940 do confirm what he remembers—we were unable to observe any field-effect results.

(I shall terminate this discussion of the copper-oxide problems, by reporting a more recent failure. In late 1972, as a possible item to report during the 25th anniversary of the transistor, I had another try with copper-oxide. Devices like those of Fig. 4(b) were made by cutting grooves in the oxide of available copper-oxide rectifiers. Like the notebook entry on the field-effect failure discussed below for 23 June 1945, the effects were inexplicably at least 1000 times smaller than expected and thus not observable. I doubt that I shall ever learn why.)

V. From 1945 Up to 17 November 1947

I was separated from solid-state physics research by World War II. I worked first on radar and later on operations research for the Navy and for the Army Air Forces. During that time, semiconductor technology markedly advanced, stimulated by the need for point-contact, or "cat's-whisker," detectors for radar. Bell Laboratories actively contributed to detector development, especially at its Holmdel location, while materials were improved by the metallurgical work at the Murray Hill Laboratory.

Silicon and germanium, both elements of the fourth column of the periodic table, became two of the technologically best-controlled semiconductors in existence. Methods of using impurities from the third and fifth columns of the periodic table to act as acceptors and donors were developed.[3] The designations p-type and n-type were in common use. Compensation of donors and acceptors was used at Murray Hill to control or to adjust resistivity and, indeed, was patented by J. H. Scaff and H. C. Theuerer. A particularly important form of high-purity, n-type germanium was developed. In point-contact rectifiers it supported high reverse voltages, also called "back" voltages, and was thus referred to as "high back-voltage germanium," a type of germanium destined to play a vital role in the discovery of the point-contact transistor.

During 1945, in anticipation of the end of World War II, Dr. Kelly took steps to return Bell Laboratories to a peace-time basis. He invited me to make several trips from my Pentagon office to Bell Laboratories. After the fall of Japan, I returned fulltime as a cosupervisor with S. O. Morgan of a solid-state physics group that included research on semiconductors, magnetic materials, and several other areas of solid-state physics.

On one of my visits, I attended a conference arranged by Dr. Kelly with Dr. Fisk also present, held at the Holmdel location. There, R. S. Ohl had made radar crystal detectors during the war. Near the end of the war Ohl had succeeded in using some of these in an amplifier. This amplifier depended upon a negative resistance effect, probably produced by what is usually called thermistor action; with the aid of this resistive effect, losses in electrical circuits could be canceled so that high amplitude oscillation could be provoked by radio waves. In fact, power gain could occur. Ohl demonstrated that amplified radio broadcasts could be heard over a small loudspeaker.

Ohl's radio set was indeed an exciting solid-state development. However, instability due to the negative resistances made it unreliable. A basic difficulty with negative resistance devices of the type used by Ohl is that they have only two terminals, so that if they give gain when incorporated in a circuit working from a low power level input to a high power level output, the circuit also has gain in the opposite direction. Thus, for high gain conditions, very marginal stability results, and small changes can produce spontaneous oscillations.

(About seven years later W. P. Mason, a Bell Laboratories colleague, and I invented a stable "dissected amplifier" by combining two-terminal negative resistors with unsymmetrical Hall-effect plates to obtain one-way gain: W. P. Mason and W. Shockley 2,775,658 issued 25 Dec 56, filed 1 Aug 52; two by W. Shockley 2,794,864 issued 4 Jun 57, filed 1 Aug 52 teaching how to make combinations that mimic vacuum-tube tetrodes and 2,777,906 issued 15 Jan 57, filed 26 Jun 53 applicable to wave guide structures. I have not kept up to date and do not know if these have any relevance to Esaki or Gunn or other negative resistance diodes.)

During one of my visits to Bell Laboratories before World War II ended, I resumed efforts to dream up a semiconductor amplifier. So far as I can now reconstruct my thinking, I did not, at that time, review the copper-oxide Schottky-barrier devices of Fig. 4 that I had thought about in 1939 and early 1940. Instead, I made use of ideas associated with the development of the technology for silicon and germanium that had occurred during the war. Fig. 5(a) shows one of my earliest postwar ideas. It consists of a series of layers in which the working current flows within the layers and the control current is applied transverse to the layers. A key advantage of this idea is that the control current does not add any net charge to the structure so that the time lag required to conduct charge along the layers is eliminated, a feature beneficial to high-frequency performance. I devised several schemes which give such control.

16 April 1945 is the date of the notebook entries of Fig. 5(b) and (c). The single layer of Fig. 5(c) contain a p-n junction. Thus, the conductivity of the exposed faces are of opposite conductivity types. Consequently, a transverse field puts opposite charges on the p-type and the n-type faces and produces depletion layers on both the surfaces. The resulting forward current across the junction will reduce the total number of carriers with resultant loss of conductivity in the plane of the layer. This idea was not pursued actively; instead, a basic field-effect structure was tried—an example of what I have described as a "try simplest cases" approach.

The "try simplest cases" field-effect concept of 1945 uses a thin layer of n-type semiconductor as shown in Fig. 6.

[3] Regarding these definitions, see footnote 2 in Section II.

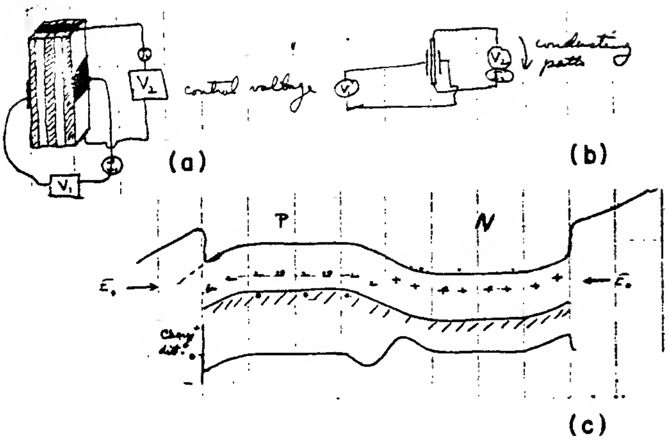


Fig. 5. Field-effect disclosures in Shockley's notebook of ideas developed after a five-year gap caused by World War II. (a) An amplifying structure and circuit, dated 13 April 1945, composed of a sequence of parallel layers. (b) A similar amplifying circuit, dated 16 April 1945, using a single layer, and (c) a proposed layer structure utilizing space-charge penetration at two surfaces without any net charge being added to the layer by virtue of the presence of a p-n junction.

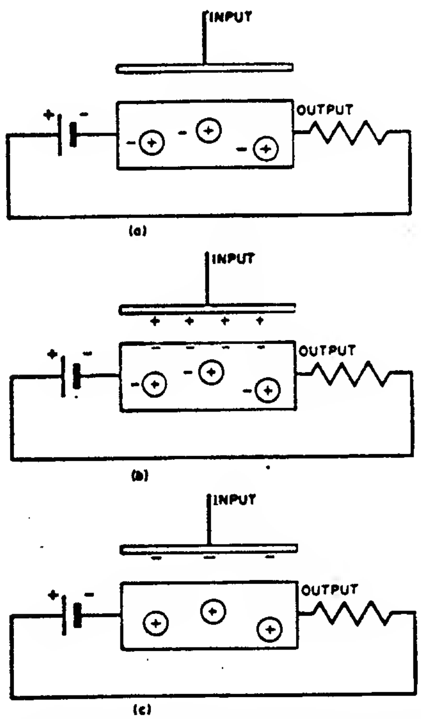


Fig. 6. The theory of a field-effect transistor using a thin layer of semiconductor: (a) The structure of the transistor with no control voltage applied. (b) The situation prevailing when a positive charge is placed on a control plate to increase the conductance of the semiconductor. (c) The situation when a negative charge is put on the capacitor plate to reduce the conductance of the semiconductor.

This layer forms one plate of a parallel plate capacitor, and the other plate is a sheet of metal. Charging the capacitor, alters the number of electrons on the semiconductor, modulating the conductance of the layer much as did the reverse bias on the copper plate in the concept of Fig. 4(b).

Thin films of silicon that had been deposited, largely by Gordon Teal, seemed ideally adapted to field-effect experiments. My calculations showed that very substantial modulation of the resistance should occur but no effect was detected.

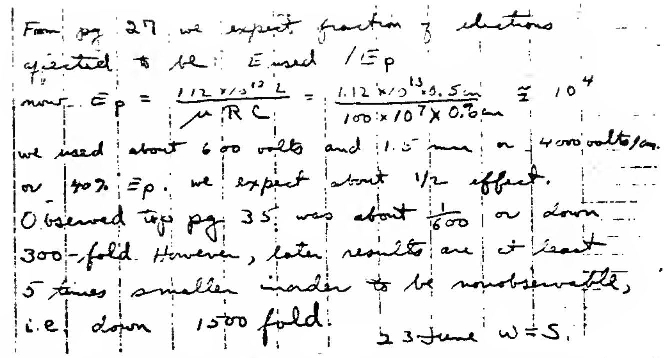

From pg 27 we expect fraction of electrons ejected to be E used / E_p
now $E_p = \frac{1.12\times10^{13} L}{\mu R C} = \frac{1.12\times10^{13}\times0.5\,cm}{100\times10^7\times0.6\,cm} \cong 10^4$
we used about 600 volts and 1.5 mm or 4000 volts/cm.
or 40% E_p. we expect about 1/2 effect.
Observed top pg 35 was about $\frac{1}{600}$ or down 300-fold. However, later results are at least 5 times smaller in order to be nonobservable, i.e. down 1500 fold.
23 June W.S.

Fig. 7. W. Shockley notebook entry of 23 June 1945 noting that the theoretically predicted effect for thin semiconductor films is so large that it should be readily observable. Actually, negligible effects were then observed indicating that the theory predicts effects at least 1500 times larger than actually occur.

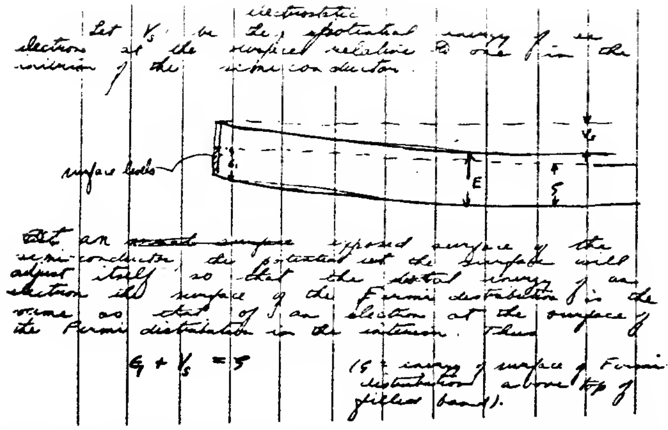
23 June 1945 was the date on which I estimated quantitatively the degree by which the effects must fall short of my theoretical expectations in order to be undetectable. Fig. 7 shows my notebook entry of that date and my calculation of the field necessary in a structure like that of Fig. 6 in order to remove all of the carriers from the specimen. I concluded from this research that the effects were at least 1500 times smaller than theoretically expected—clear evidence that something was the matter and that we did not have an adequate scientific understanding of certain basic phenomena that were connected with field-effect amplifying devices.

I brought my calculations to the attention of John Bardeen who had joined Bell Laboratories at my invitation in 1945. He confirmed, about two weeks later, the accuracy of my calculations and was equally puzzled by the discord between theory and experiment.

19 March 1946 was the date at which Bardeen converted the puzzling field-effect failure into a creative step. He wrote in his notebook one of the most significant research ideas of the semiconductor program, shown in Fig. 8. Fig. 9 presents this idea as it applies to the failure of the field-effect for the structure of Fig. 6. Bardeen's explanation for the field-effect failure was this: Fig. 4 was in error by suggesting that electrons drawn to the surface of the semiconductor, by charging it negatively were as free to move as were electrons in the interior. Instead, these added electrons were trapped, immobile, in surface states, as represented in Fig. 9. In effect, the surface states blocked the external field at the surface and, thereby, shielded the interior of the semiconductor from the influence of the positively charged control plate.

Bardeen found that his surface states explained not only the field-effect failure but also a number of previously perplexing observations about semiconductor surfaces. These other phenomena were important in respect to the research program in general but had little relevance on the path to the junction transistor—with one very important exception—the one associated with the date of 17 November 1947 discussed in Section VI.

My interests in p-n junctions were apparently nearly negligible in 1946 but became active again in 1947.


Let V_s be the electrostatic potential energy of an electron at the surface relative to that in the interior of the semi-conductor.

surface levels

At an ~~ideal~~ surface exposed surface of the semi-conductor the potential at the surface will adjust itself so that the total charge of an electron at surface of the Fermi distribution in the volume as that of an electron at the surface of the Fermi distribution in the interior. Then

$\zeta_s + V_s = \zeta$ (ζ_s = energy of surface of Fermi distribution above top of filled band).

Fig. 8. John Bardeen's notebook entry of 19 March 1946 presenting the first record of his proposal of surface states to explain the failure reported in Fig. 7.

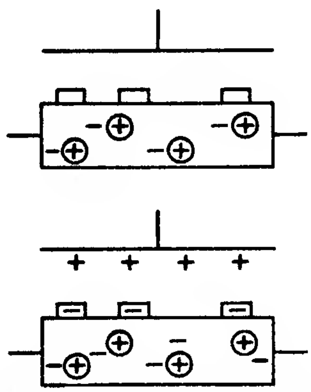


Fig. 9. Bardeen's explanation of the failure of the field-effect structure through the presence of surface states. (a) The structure with no applied voltage. (b) The capture in surface states of the electrons induced by the surface charge.

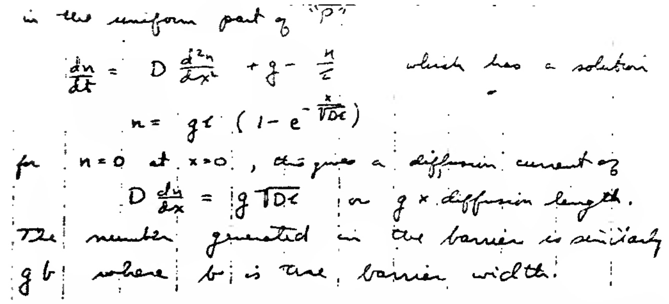
12 March 1947, the date when I considered a p-n junction from the point of view of relating its electrical properties to fundamental physical mechanisms. The diagram, not reproduced here, shows an energy band diagram for a p-n junction at thermal equilibrium and makes this following observation:

> The resistance of this junction involves the recombination of electrons with holes. For zero bias, the product of hole times electron concentration is uniform over the junction and involves the energy gap in $\exp(-E/kT)$. Hence the temperature coefficient of resistance should give the energy gap. A capacity measurement should give the barrier spacing. I suspect that knowing these two, the absolute value of the resistance will give a value for the recombination cross section of holes and electrons.

The emphasis here is on the relationship of electrical characteristics to fundamental physical mechanisms. There is no concern with p-n junctions as elements of useful devices.

4 April 1947. Here I did discuss a structure containing alternating n- and p-type layers. It came very close, but it did not define injection. This notebook entry reads as follows:

> *Lightning Arrestor*
>
> Evaporate alternate layers of N & P silicon with such donator and acceptor densities and layer thicknesses as to produce high resistance. When high field strength is applied, both electrons and holes can be pulled over the potential maxima and minima and higher conductivity will result. 4 Apr 47, W. Shockley


in the uniform part of "P"

$\frac{dn}{dt} = D\frac{d^2n}{dx^2} + g - \frac{n}{\tau}$ which has a solution

$n = g\tau(1 - e^{-\frac{x}{\sqrt{D\tau}}})$

for n = 0 at x = 0, this gives a diffusion current of

$D\frac{dn}{dx} = g\sqrt{D\tau}$ or g × diffusion length.

The number generated in the barrier is similarly gb where b is the barrier width.

Fig. 10. W. Shockley notebook entry of 24 April 1947 consisting of what is apparently the first record of diffusion theory for minority carriers in a p-type layer adjacent to a reverse-biased p-n junction.

Two months later, on 10 June 1947, this entry was endorsed as "witnessed and understood" by J. Bardeen and Walter H. Brattain. "Why two months later?" I wondered. "What happened on 10 June 1947?" Research on other notebooks gave the answer: On the same date, I had witnessed a slightly related disclosure by Walter Brattain which must have reminded me of my 4 April 1947 lightning arrestor idea and led to my asking that its disclosure be witnessed.

The 4 April 1947 disclosure has features like those of the junction transistor. In the lightning arrestor, reverse-biased p-n junctions collect minority carriers passing through intermediate layers. But these concepts did not occur in a form that crystallized into the pattern needed for the concept of the junction transistor.

My interest in lightning arrestors had been stimulated by a post World War II suggestion, like M. J. Kelly's of about eight years before, by O. E. Buckley, then President of Bell Laboratories.

24 April 1947 was a milestone date in my notebook. I then considered current–voltage characteristics and basic mechanisms in p-n junctions. The key features appear on page 77 of my notebook and are reprinted in Fig. 10. This is the first entry that I found analyzing the behavior of minority electrons in a uniform p-layer. Incidentally, this may be one of the earliest records of the now generally accepted usage of "n" and "p" for electron and hole densities. Customary usages in 1947 were "n_e" for "excess electrons" or electrons in the "empty band" and "n_h" for "holes in the filled band."

The focus of interest, when I did the analysis of Fig. 10, was on the physical phenomena and not on device applications. My interest was in using electrical measurements to obtain fundamental scientific information about basic processes.

The conspicuous shortcoming in Fig. 10 is its limitation to the case of large reverse bias. This limitation does simplify the boundary condition at the edge of the space-charge layer to zero minority-carrier density, but it did prevent discovering minority-carrier injection and its exponential increase with forward bias.

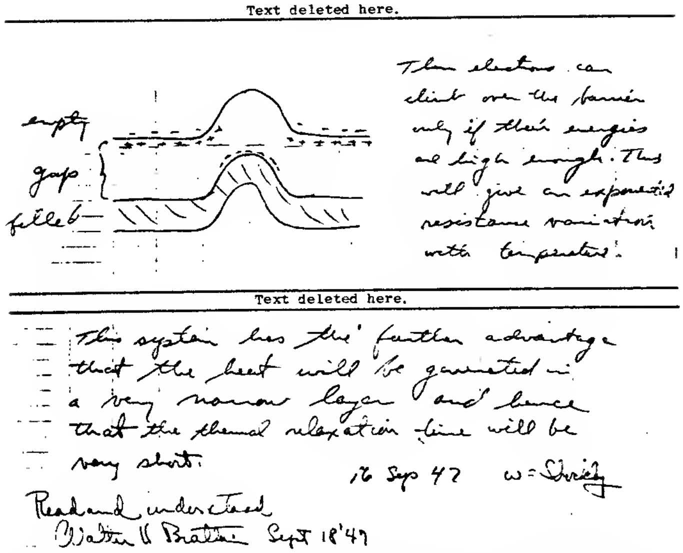


Fig. 11. W. Shockley 16 September 1947 notebook entry witnessed by W. H. Brattain two days later showing a three-layer n-p-n structure intended as a high-speed thermistor involving minority carriers passing through the intermediate layer. This is an example of a structure which might, but did not, lead to the conception of minority carrier injection through the base layer of a transistor. (Note the then conventional use of "filled" and "empty" for what are now "valence" and "conduction"—a convention introduced, I believe, in my book *Electrons and Holes in Semiconductors* [12].)

On the other hand, I note that I did consider the generation of carriers in the space-charge layer, a phenomenon that gives rise to significant currents in reverse-biased silicon p-n junctions, a subject that I analyzed with T. Sah and R. N. Noyce about ten years later.

19 May 1947 shows another entry related to reverse currents in p-n junctions. It reads:

REVERSE CURRENT IN N-P JUNCTIONS.

These currents may be of great theoretical interest because a failure to saturate may represent transitions from the full to empty bands induced by the field. Such transitions have been discussed by Zener and others from a theoretical viewpoint and the N-P barrier may afford one of the first opportunities to study them experimentally. The adverse effects of such transitions may be reduced by increasing the barrier thickness since this will reduce the field strength. The barrier thickness will be inversely proportional to impurity concentration gradient and can thus be affected by annealing and by controlling impurity contents and gradients. 19 May 47, W = Shockley.

(About four years later, I combined my interests in lightning arrestors and Zener currents in the Zener-diode patent: W. Shockley 2,714,702 issued 2 Aug 55 filed 16 Feb 51. For references, see [6] for MacAfee *et al.* and for McKay on avalanche interpretation.)

16 September 1947. On this date I returned again to thinking about n-p-n structures as shown in Fig. 11. For this structure, I once more considered a flow of minority carriers through an intermediate layer in an n-p-n structure. My idea was to build a negative resistance device having high frequency response through thermistor action, the same principle proposed above to explain the amplifying devices made by R. S. Ohl at the Holmdel Laboratory. What I proposed, as shown in Fig. 11, was that raising the temperature would increase exponentially the conductivity of electrons through the intermediate p-type layer. I also noted that since the heat is generated in a very thin layer, this structure will have a short thermal relaxation time and, therefore, might follow high frequencies. This particular idea was read and understood by Walter H. Brattain, who signed the disclosure on 18 September 1947.

But this n-p-n disclosure did not quite achieve the concept of a current of injected carriers into the middle layer that increases exponentially, not with temperature, but with forward voltage across a p-n junction. Similar limitations applied to my other background ideas, discussed above in this section, about p-n junctions.

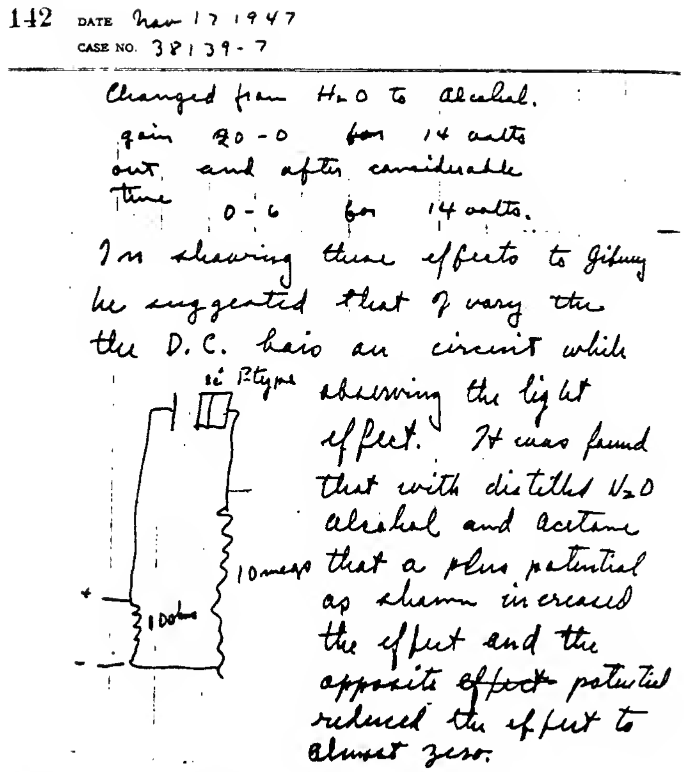
VI. The "Will to Think" and the Birth of the Point-Contact Transistor

In the Solid-State Physics Department during the Fall of 1947, one of the most active research topics in the transistor group—really the pretransistor group then—was Bardeen's surface states theory. According to Bardeen, the surface of the semiconductor might be so charged as to repel majority carriers. In the case, to be discussed below, of a p-type silicon specimen, a positive surface-state charge would repel holes, the majority carriers in the specimen. If photons of light fall on the semiconductor and generate hole–electron pairs, then the field just below the surface will separate them, electrons going to the surface and holes to the interior. This separation causes the surface potential to become more negative.

Such effects were being actively studied by Brattain. As he and Bardeen had found, this photovoltaic effect was one of the most effective methods of demonstrating the existence of Bardeen's surface states and of studying their properties. Brattain used a chopped light source that produced an ac voltage on the silicon surface, detected and measured by using a metal screen parallel to the surface.

Monday, 17 November 1947. This notebook entry by Brattain, reprinted in Fig. 12, reports the results of a suggestion made by Robert B. Gibney, a physical chemist whom I had recruited from the chemistry department soon after the war to add his training in physical chemistry and to round out the scientific expertise of the semiconductor group. (Climatic reasons led him to find and to accept a position at Los Alamos in favor of New Jersey and we lost him in early 1948, but only after he had contributed greatly to the point-contact transistor program.)

As recorded by Brattain in Fig. 12, Gibney made the key suggestion that voltage be applied between the metal plate and the semiconductor while both were immersed in an electrolyte. By this means a strong electric field was generated perpendicular to the semiconductor surface. The effects observed by Brattain showed that, when the polarity of the applied field was such as to make the silicon surface more positive, there was an increase in the negative photovoltaic impulse produced on the surface by the



142 DATE Nov 17 1947
CASE NO. 38139-7

Changed from H_2O to alcohol.
gain 20-0 for 14 volts
out, and after considerable
time 0-6 for 14 volts.
I'm showing these effects to Gibney
he suggested that I vary the
the D.C. bias on circuit while
observing the light
effect. It was found
that with distilled H_2O
alcohol and acetone
that a plus potential
as shown increased
the effect and the
opposite ~~effect~~ potential
reduced the effect to
almost zero.

Fig. 12. Brattain's notebook entry of 17 November 1947 that motivated the "will to think" and initiated the "magic month" culminating in the invention of the point-contact transistor on 16 December 1947.

turn-on of the light. This new phenomenon was simply explained by assuming that the internal field inside the semiconductor had been increased so as to repel the majority holes further from the surface. This larger field drew photoelectrons more effectively to the surface and, thereby, increased the effect produced when the illumination increased. This is the significance of Brattain's words in Fig. 12 "that a plus potential as shown increased the effect." Similar reasoning explains why "the opposite potential reduced the effect to almost zero."

This new finding was electrifying. At long last, Brattain and Gibney had overcome the blocking effect of the surface states—the practical problem that had for so long caused the failure of our field-effect experiments.

"Accident favors the prepared mind"—in this case, minds prepared to try to make field-effect amplifiers—explains the stepped-up use of notebook pages in Fig. 3. The new experimental conditions set up to study surface states and the resulting new facts about penetration of the field into the surface motivated the will to think of ways to achieve the practical result of a semiconductor amplifier. Within a week, Bardeen, Brattain, and Gibney had conceived two of the five device inventions that were filed for patenting before the public announcement of the transistor on 30 June 1948.

Thursday, 20 November 1947. Only 3 days after the new facts were observed, Brattain and Gibney wrote a disclosure of the concept of patent 2,524,034 (filed 26 February 1948; issued 3 October 1950). They proposed amplifiers using the field effect with an electrolyte to obtain the desired high electric fields. The actual structure which worked a few days later is presented in Brattain's notebook entry of 4 December 1947 discussed below. They also suggested that solids could be used in the concluding sentence of their disclosure:

> It is of course evident that the liquid dielectric could be replaced by a solid dielectric if one can be found having the proper ionic mobility to form such a dipole layer at the surface of the semiconductor.

On the same day that it was written, this disclosure was witnessed by John Bardeen and Hilbert R. Moore, an electrical engineer who was also a member of the semiconductor group.

During the week starting on 17 November 1947, the time lag between the writing of a disclosure and the obtaining of the signature of a witness was not two months, as it had been for my 1939 conception of Fig. 4. It was not even as much as two weeks, such as had elapsed in 1945 for Bardeen's check of my field-effect calculations of Fig. 7.

The stepped-up tempo of November 1947 compared to earlier times is easy to explain. At those earlier dates, we were doubtful that thinking would produce worthwhile action. The breakthrough observation of 17 November that surface-states could be overcome stimulated *the will to think*—and to act—in minds conditioned to search for semiconductor amplifiers.

Sunday, 23 November 1947. This stimulus is dramatically illustrated by Bardeen's notebook disclosure of this date of a field-effect concept that led to his patent (2,524,033 filed 26 February 1948; issued 3 October 1950)—one of the five patent applications for transistor devices filed before the public announcement of the transistor. Bardeen introduced his disclosure by referring to Brattain and Gibney's observation of the "Shockley effect" and then writing: "It occurred to the writer that the effect might be observed in the thin n-type layer on the surface of a block of p-type Si."

I reprinted the diagram of Bardeen's disclosure as Fig. 20 of [3] and described it as "an insulated-gate field-effect transistor with an inversion layer channel"—in other words, an anticipation of an IGFET. I was incorrect in this interpretation.

What Bardeen disclosed was not the type of inversion-layer channel, dynamically induced by gate voltage, that is basic to the IGFET's that were to be developed more than ten years later. Bardeen's channel was "a surface layer of the order of 10^{-4} cm in thickness, separated from the body by a barrier [a p-n junction]," in the words of his patent. His patent cited an invention (another of the five patents: Robert B. Gibney 2,560,792 filed 26 February 1948; issued 17 July 1951) for preparing n-type germanium so as to form "a thin surface layer of p-type material containing fixed negative charges and mobile positive charges."

Bardeen's invention was a creative failure. More than ten years were needed before concepts and technology led to good IGFET's. In a sense, it was like my Schottky-gate idea of 1939 by being ahead of its time. However, Bardeen's

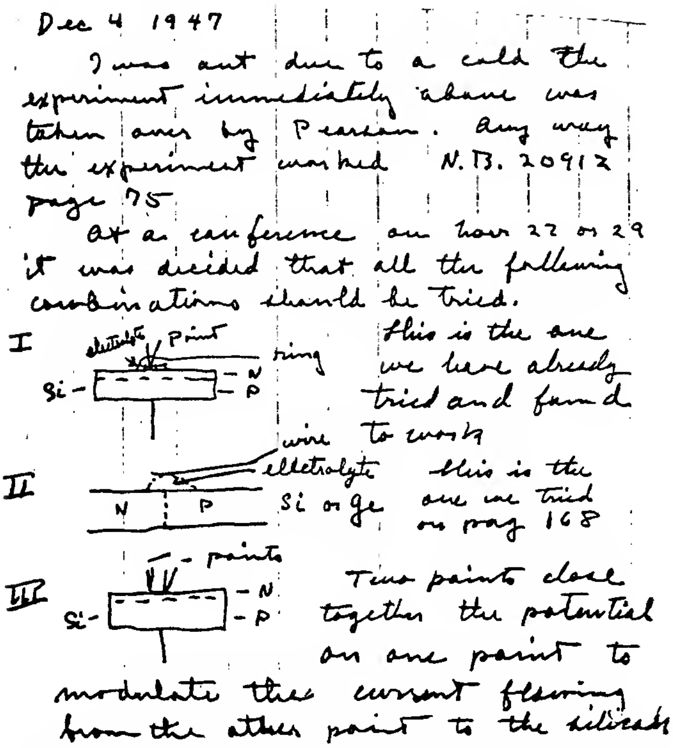


Fig. 13. Three key experiments recorded in Brattain's notebook of 4 December 1947.

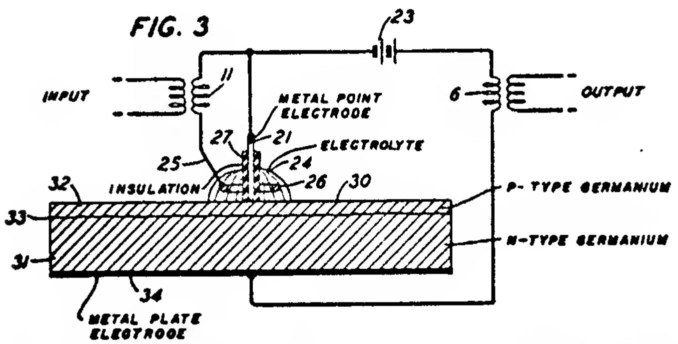


Fig. 14. A drawing from the Brattain–Gibney patent corresponding to Experiment I of 4 Dec. 47. The p-type layer on n-type germanium is replaced by an n-type layer on p-type silicon in a similar drawing in the Bardeen patent. In both cases, a preferred electrolyte is the so-called glycol-borate "Gu."

thinking about it played a key role in reaching the point-contact transistor. I discuss it at length to illustrate how the breakthrough observation of 17 November 1947 stimulated the will to think. Prior to that observation, although he was familiar with the concepts of both surface layers and also of field-effect amplification, Bardeen had not combined the two to arrive at the concept of his 23 November 1947 disclosure.

Thursday, 4 December 1947. Fig. 13 shows Brattain's summary of several experiments performed during the preceding week. The experiments designated I and II had both been successfully tried. Experiment I supported two of the patents cited above: Brattain–Gibney and Bardeen. The patent drawing of Fig. 14 of Brattain–Gibney, like a corresponding drawing from Bardeen, shows that the mechanism consists of modulating the resistance of a thin surface layer with voltage applied to an electrolyte. A point contact, insulated from the electrolyte, carries current to the layer. Experiment III anticipates the structure of the point-contact transistor.

Experiment II lies on the path to the concept of the junction transistor. In doing research for this historical note, I referred to Pearson's notebook to verify Brattain's statement in Fig. 13 that, as a result of Pearson's efforts, my suggestion (recorded in Fig. 13 as appearing on Brattain's p. 168) had worked. It had. When I analyzed Pearson's data, I found that they predicted power gain in the proper circuit.

I found something else that I had not remembered at all. Pearson had done many field-effect experiments in the past and had observed feeble effects at liquid nitrogen temperatures. But after the advent of the electrolytes, he had been far more successful. The favorite electrolyte, glycol borate, called "Gu," was essential.

"Gu" was obtained by extracting it from electrolytic capacitors by using a vice, a hammer, and a nail.

What I had forgotten was in Pearson's notebook on several pages following the results of experiment II. Pearson described how Gibney had prepared thin film specimens by evaporating high back-voltage germanium onto hot quartz plates. These improved versions of the field-effect arrangement of Fig. 6 worked. Pearson could reduce the resistance by 28.4 percent or increase it by 34.8 percent. About these results, Pearson wrote in his notebook:

> I will make a complete curve on this next. We then want to do the same curve on n-type which should decrease with plus voltage and increase with minus. This is a moral victory because it is a positive result on the field effect which we have been looking for so long. It is the use of the Gu which enables us to get the high fields necessary to successfully perform the experiment as well as using the crystalline film.

The experiment shown in II of Fig. 13 was not merely a step on the path to the junction transist but actually a feature of the issued patent. It appears in the patent drawings of Fig. 18 as discussed in Section VII.

Monday, 8 December 1947. This was the date that initiated my own stepped-up tempo of notebook pages per month. It seems obvious from the collection of items in Fig. 15 that I had been infected by the will to think about new kinds of semiconductor amplifiers, especially those using p-n junctions. The entries of Fig. 15 commence with the disclosure of a junction field-effect amplifier. (Research on the Patent Department files show that I did not recall this disclosure [I referred to later conceptions and an achievement by Gerald Pearson] when preparing a history of invention for my junction field-effect transistor patent 2,774,970 issued 8 May 1956, filed 24 August 1951, that introduced source-gate-drain and pinch-off. See [16] for creative-failure aspects and reference to first 1952 unipolar-transitor article.)

Fig. 15 contains an important stepping stone on the path to the concept of the junction transistor. This is concept (2) discussed with Fig. 1, the use of a reverse-biased p-n

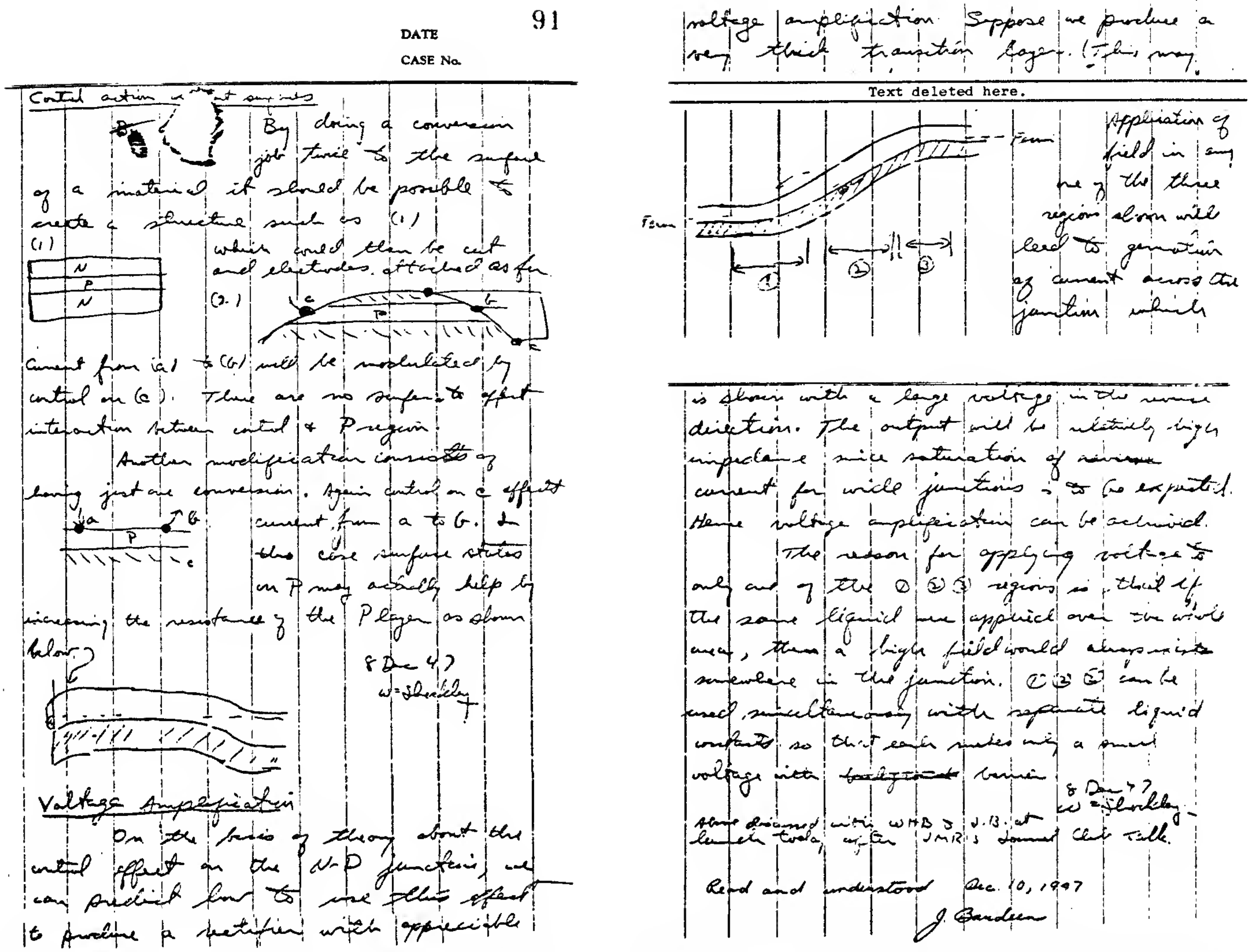


Fig. 15. Several disclosures of W. Shockley as a notebook entry on 8 December 1947. One of these figures shows the first example of an n-p-n field-effect structure but not a junction transistor. Another item includes a proposal for voltage gain by using a wide p-n junction operated in the reverse direction at high reverse voltage.

junction to obtain voltage gain. The proposed mechanism was based on the success of the drop of electrolyte over the junction in Brattain's experiment II of Fig. 13. Placing drops of electrolyte at localized intermediate points on the junction might control reverse current using relatively small voltages applied to the electrolyte and thus produce large voltage changes across the p-n junction.

The actual working mechanism of experiment II was probably the formation of surface channels near the junction. These would then increase the reverse current. (About 17 years later, I conceived a somewhat similar device based on sound principles. This "surface-controlled avalanche transistor" (SCAT) is covered by patent 3,339,086, filed 11 June 1964 and issued 29 August 1967.)

However, this half-baked reverse-biased p-n junction idea of mine for an amplifier did actually pay off in terms of progress toward achieving the point-contact transistor. This progress followed a lunch time discussion described next.

Monday, 8 December 1947. On this date significant progress was made towards the point-contact transistor. Walter Brattain's notebook reports observation of a voltage gain of 2 and a power gain of 330 in an entry signed by both Bardeen and Brattain. The amplifying device which they used consisted of the drop of electrolyte and point-contact structure of experiment I of Fig. 13 and Fig. 14. However, a key new feature was the material used—high back-voltage n-type germanium—later a central feature in achieving the voltage gain of the point-contact transistor. Brattain's entry describes how this came about as follows:

> Note the above result or experiment was started as result of a luncheon discussion with Shockley and Bardeen and the final suggestion by Bardeen that voltage amplification could be obtained if the above experiment was performed on high back-voltage Ge . . . .

At that luncheon, I had discussed the concept of my notebook entry using a reverse-biased junction for voltage gain. Bardeen had converted this idea to a usable form with reverse bias on high back-voltage germanium. The circumstance of this discussion presents an important illustration of the mixture of cooperation and competition that characterized the interactions within the semiconductor group.

The circumstances of Fig. 15 show creative-failure methodology in action. My reverse biased p-n junction concept was basically a failure. But as modified by Bardeen, it became a significant forward step to the point-contact transistor. The experience with it also helped to prepare my mind to take advantage of the subsequent opportunity, discussed in the next section, to conceive of the junction transistor. The authenticity of this reconstruction of the events of 8 December 1947 is supported by the notebook records quoted above and shown in Fig. 15 for 8 and 10 December. A similar reconstruction dependent solely on memory would be far less reliable.

16 December 1947. An entry in Brattain's notebook clearly shows that on this date the point-contact transistor "was born." In his Nobel lecture [5] in 1956, John Bardeen described the sequence of events between 8 and 16 December:

> It was next decided to try a similar arrangement with a block of n-type germanium. Although we had no prior knowledge of a p-type inversion layer on the surface, the experiments showed definitely that a large part of the reverse current consisted of holes flowing in an inversion layer near the surface. A positive change in voltage on the probe [the ring in the electrolyte of Figs. 13 and 14] decreased the reverse current. Considerable voltage as well as current and power amplification was observed.
>
> Because of the long time constants of the electrolyte used, amplification was obtained only at very low frequencies. We next tried to replace the electrolyte by a metal control electrode insulated from the surface by either a thin oxide layer or by a rectifying contact. A surface was prepared by Gibney by anodizing the surface and then evaporating several gold spots on it. Although none made the desired high resistance contact to the block, we decided to see what effects would be obtained. A point-contact was placed very close to one of the spots and biased in the reverse direction [In [5] Bardeen here refers to a figure]. A small effect on the reverse current was observed when the spot was biased positively but of the *opposite* direction to that observed with the electrolyte. An increase in positive bias *increased* rather than decreased the reverse current to the point contact. The effect was large enough to give some voltage, but no power amplification. This experiment suggested that holes were flowing into the germanium surface from the gold spot, and that the holes introduced in this way flowed into the point contact to enhance the reverse current. This was the first indication of the transistor effect.
>
> It was estimated that power amplification could be obtained if the metal contacts were spaced at distances of the order of 0.005 cm. In the first attempt, which was successful, contacts were made by evaporating gold on a wedge and then separating the gold at the point of the wedge with a razor blade to make two closely spaced contacts. After further experimentation, it appeared that the easiest way to achieve the desired close separation was to use two appropriately shaped point-contacts placed very close together. Success was achieved in the first trial; the point-contact transistor was born [5].

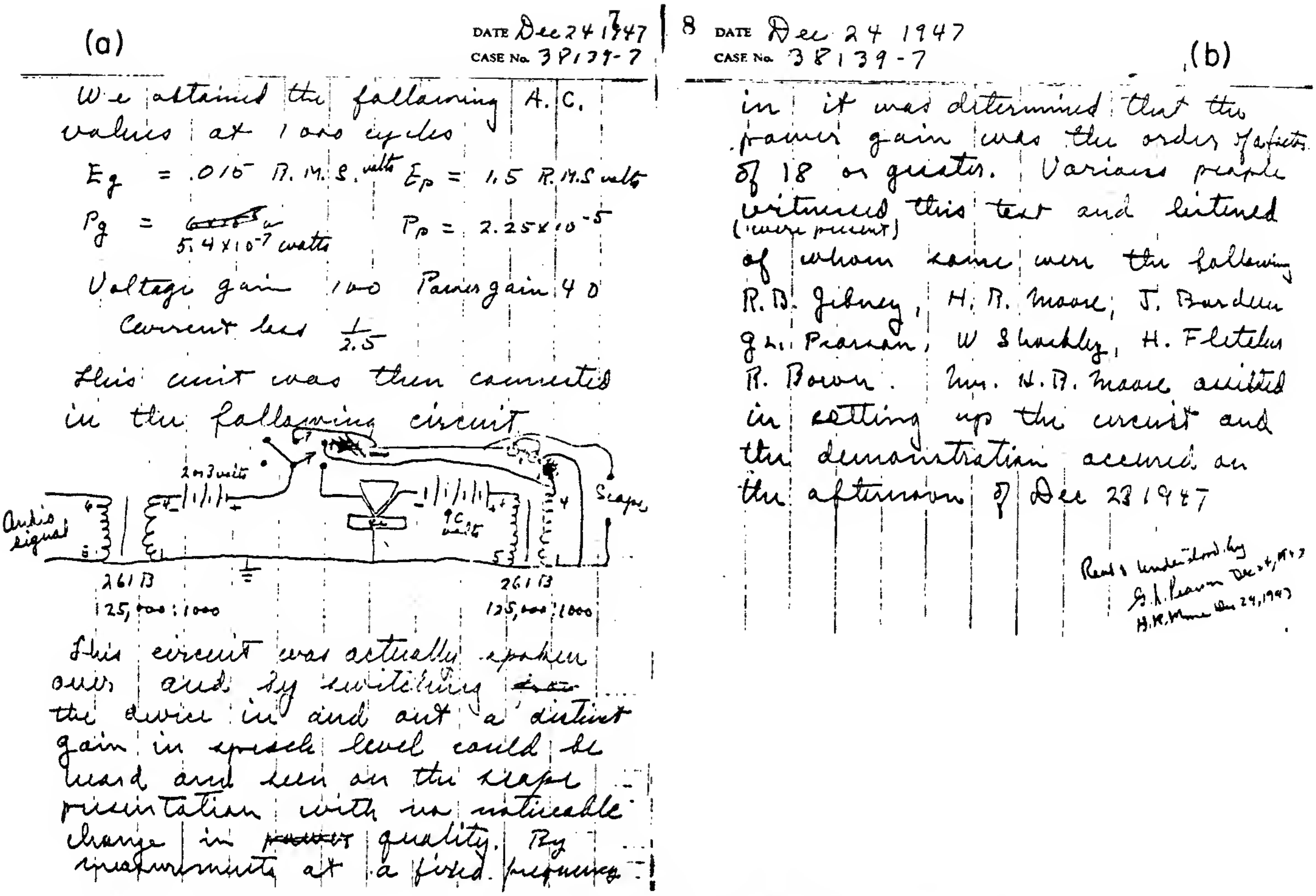
Fig. 16 shows the original model described by Bardeen.

Tuesday, 23 December 1947, was the date of the private demonstration for executives that is described in Brattain's famous notebook entry of Christmas Eve, 1947, Fig. 17. Although this date of 23 December 1947 has been publicly accepted in some instances as the date for the birth of the transistor, my research on laboratory notebooks confirms Bardeen's statement, quoted above, that the first try with the wedge structure of Fig. 16 succeeded and the point-contact transistor was born. The date of that experiment was 16 December 1947.

Fig. 16. The original point-contact transistor structure comprising the plate of n-type germanium and two line-contacts of gold supported on a plastic wedge. (The name "base," which arose from this structure, does not have functional significance as do "emitter" and "collector.")

The period of 29 days from the breakthrough observation of 17 November to the point-contact amplification of 16 December thus includes so many creative contributions and started and finished with such significant events as to deserve the title "the magic month."

There is an element of obvious—and entertaining—naiveté about accepting the day before Christmas Eve of 1947 as the date of the birth of the transistor. The demonstration of 23 December was attended by Harvey Fletcher, the Director of Physical Research, to whom I reported, and Ralph Bown, the Director of Research, to whom Harvey Fletcher reported. To believe that such a demonstration occurred on the first day that a transistor had worked—the schedules of these executives—the reliability of the demonstration equipment—a decision by the research scientists that they were sure—what better definition of naiveté? But this is an example of creative-failure methodology in public relations. The 24 December 1947 notebook entry is dramatic; hearing speech amplified by the transistor was in the tradition of Alexander Graham Bell's famous "Mr. Watson, come here, I want you!" (Unfortunately, there is no record that these words were transmitted through a transistor during the 23 December 1947 demonstration.) The failure to correctly date the birth of the transistor had pleasant consequences. It caused anniversaries of the transistor to be held in a warm Christmas Eve atmosphere.


(a)

DATE Dec 24 1947
CASE No. 38139-7

We obtained the following A.C. values at 1000 cycles

Eg = .015 R.M.S. volts Ep = 1.5 R.M.S volts

Pg = 5.4 x 10-7 watts Pp = 2.25 x 10-5

Voltage gain 100 Power gain 40

Current loss 1/2.5

This unit was then connected in the following circuit

This circuit was actually spoken over and by switching the device in and out a distinct gain in speech level could be heard and seen on the scope presentation with no noticeable change in ~~power~~ quality. By measurements at a fixed frequency

(b)

8 DATE Dec 24 1947
CASE No. 38139-7

in it was determined that the power gain was the order of a factor of 18 or greater. Various people witnessed this test and listened (were present) of whom some were the following R.B. Gibney, H.R. Moore, J. Bardeen, G.L. Pearson, W Shockley, H. Fletcher R. Bown. Mr. H.R. Moore assisted in setting up the circuit and the demonstration occurred on the afternoon of Dec 23 1947

Read & understood by G.L. Pearson Dec 24, 1947 H.R. Moore Dec 24, 1947

Fig. 17. Notebook entry by W. H. Brattain of 24 December 1947 describing the point-contact transistor demonstration of 23 December 1947 at Bell Laboratories.

There is no record that Brattain's and Bardeen's experiments of 16–22 December were formally witnessed by others—an important factor of "reduction to practice." The 24 December notebook entry was witnessed by others who confirmed the recorded facts about the 23 December demonstration. So the date of 23 December does appear on the records of the Patent Department as the date of "reduction to practice." A test of reduction to practice in patent law is whether a businessman would be prepared to invest in development. There was, of course, no doubt about this at Bell Laboratories—or was there on 23 December?

I have a clear recollection that Harvey Fletcher did raise a significant question to this effect: "How do you know you really have amplification in the telephone conversation demonstration? It may be simply matching of impedances? Making an oscillator would be a valuable confirmation." Actually the input and output voltage measurements had already clearly shown true power gain. But it is noteworthy that Brattain's 24 December notebook entry continues after the report of the 23 December demonstration to record that an oscillator was constructed on 24 December and did, indeed, oscillate.

The identification of the exact date of the "invention" of the transistor is a legal matter. If "conception" is taken as the definition of "invention," then 15 December is confirmed both by my notebook research and by the Patent Department's records—the first observation of amplification was 1 day later. "Reduction to practice" in a legally-sound form was 23 December. In any event, it was an exciting start for a four-day Christmas weekend.

VII. The Conception of the Junction Transistor

The birth of the point-contact transistor was a magnificent Christmas present for the group as a whole. I shared in the rejoicing. But my emotions were somewhat conflicted. My elation with the group's success was tempered by not being one of the inventors. I experienced some frustration that my personal efforts, started more than eight years before, had not resulted in a significant inventive contribution of my own. In response to this frustration, for the next five years, I did my best to put the Labs—and myself—in the lead for transistor patents. (I had some success; most of my 90-odd issued U.S. patents relate to the transistor.) The efforts that I made with this motivation account for much of the peak of notebook entries shown on Fig. 3 for the month from 24 December 1947 to 24 January 1948. This peak includes the four or five pages that cover the disclosure of the concept of the junction transistor.

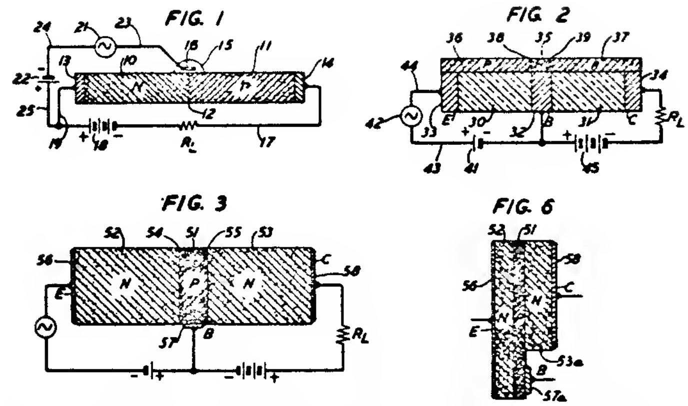


Fig. 18. Figures from the issued junction transistor patent. (2,569,347 issued 25 Sep 51 to William (Bradford) Shockley filed 26 Jun 48) The figures are discussed in detail in the text.

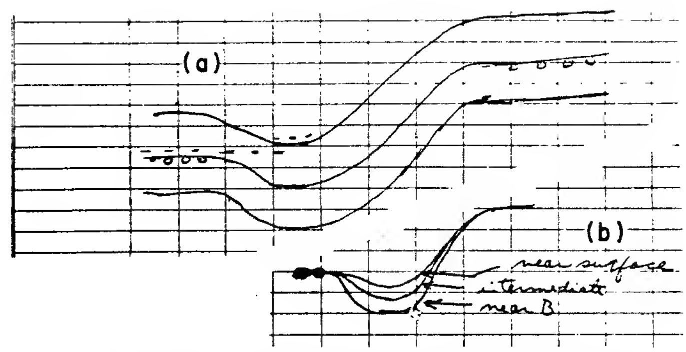


Fig. 19. A disclosure of 31 December 1947 in Shockley's notebook that should have, in spite of its practical shortcomings, suggested that minority carrier injection would be important in making transistors. (See text for the relevant details.)

The path to the junction transistor was far from direct. Perspective about at least two digressions can be obtained from the issued patent. In that patent, R. J. Guenther, the patent attorney and later the head of Bell Laboratories' Patent Department, skillfully preserved in the patent application itself the merit that any of these digressions may have had. This is illustrated by Fig. 18, a reprint of some drawings from the issued patent.

One important digression appears as "1" on Fig. 18. It is the "Gu-across-junction" discussed in Section V as experiment II of Brattain's entry of 4 December 1947, Fig. 13. When the junction-transistor patent application was filed on 26 June 1948, less than a week before the public announcement, this structure was the only one that had actually amplified. Indeed, it might have been a strong defense if the teachings of the patent had been contested. One of the claims in the issued patent would have been completely satisfied by "1" of Fig. 18 but that claim would also have been satisfied by a three layer n-p-n like that of Fig. 1. (See [3] for a detailed discussion of this feature of claim 29.)

On Fig. 18, "2" was conceived on 31 December 1947 and is the digression that led me so close to the true junction transistor that it is hard to see why it took me more than three weeks to complete the concept.

Again on Fig. 18, "6" is a draftsman's version the structure of the disclosure of 23 January 1948, the conception of the junction transistor, discussed below as Fig. 20.

The neatest structure of Fig. 18 is "3." This structure was a logical outgrowth of "6" and is what I had in mind in the photograph of Fig. 1. In fact, the structure shown as number "3" anticipates that of the first microwatt transistor that was photographed almost three years later and appears below in Fig. 25.

31 December 1947. On New Year's Eve I was alone in Chicago between two meetings that came so close together that a return to New Jersey seemed impractical. I used this opportunity of uninterrupted time to invent new semiconductor amplifying principles. In two days I wrote enough to fill a bit more than 19 notebook pages. My notebook was at Bell Laboratories and I used a pad of paper and mailed the disclosures back to my cosupervisor, S. O. Morgan, who witnessed them and asked Bardeen to do the same. Later these pages were rubber-cemented into my notebook where they remained available for study while writing this article.

The ideas related to "2" of Fig. 18 covered three of the five pages that I wrote on New Year's Eve. Fourteen more were written on New Year's Day of 1948. The structure of number "2" in Fig. 18 has two p-type regions separated by a strip of n-type. The n-type strip is formed by heating a thin layer of germanium lying upon a plane formed by two ceramic insulators separated by a thin slab of antimony-bearing alloy. The antimony diffuses into the germanium converting a strip into n-type. I shall not consider the naiveté of the fabrication scheme. I shall instead focus on a blind spot in the amplification concept. These notebook entries reveal that I missed an obvious opportunity. I failed to recognize the possibility of minority carrier injection into a base layer. I did so even while considering a device containing a base "strip" that led to an energy-band diagram that was almost indistinguishable from that for a true junction transistor. This may be seen from the reproduction of that lost "lost-opportunity" energy-band diagram reproduced from my notebook disclosure in Fig. 19(a).

What Fig. 19(a) represents is the variation along a line extending from the p-type emitter through the n-type base strip to the p-type collector. It looks precisely like the diagram for a true junction transistor. But that wasn't the idea at all. The base was not a "layer" but a "strip" formed by diffusion from *B* in "2" of Fig. 18.

How the amplification concept of this disclosure differed from that of a true junction transistor is represented in Fig. 19(b). The antimony was supposed to convert the germanium to strong n-type near the antimony-bearing center contact, referred to as *B* in Fig. 19(b). The conversion hardly occurs at all near the surface. Thus the barrier for hole flow was low near the surface and high near *B*, as indicated in the diagram. By applying negative voltage to *B*, the barrier could be lowered so that holes could easily flow over it to reach the reverse-biased collector. The disclosure also suggests using the structure as a junction field-effect

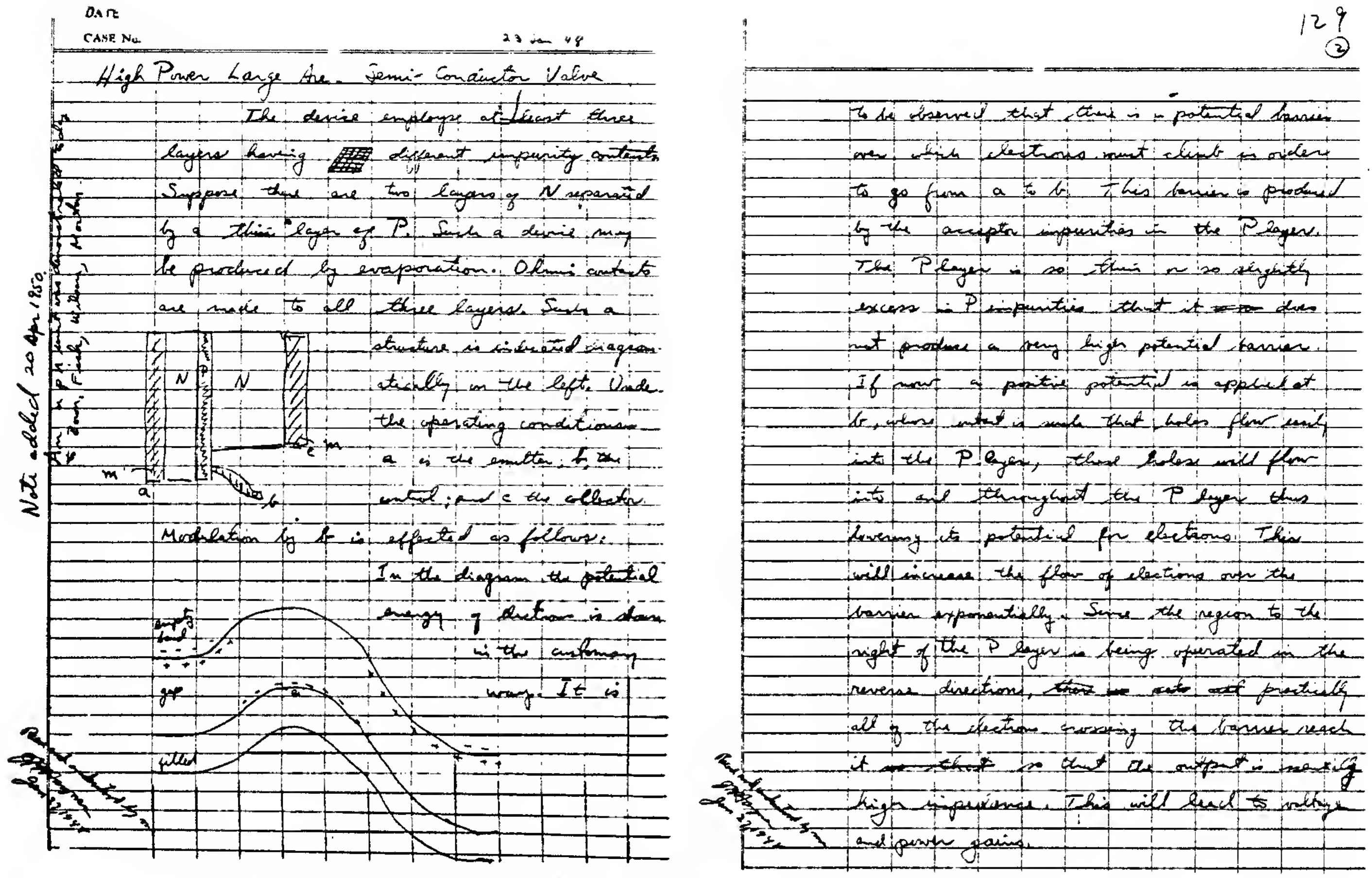


Fig. 20. The first two of five pages dated 23 January 1948 from Shockley's laboratory notebook containing the record of the combination of the three concepts discussed in Section II, the end of the path to the conception of the junction transistor.

transistor with a superficial p-type layer channel lying over the n-type strip gate.

What is conspicuously *lacking* is any suggestion of the possibility that holes might be injected into the n-type material of the strip itself, thereby, becoming minority carriers in the presence of electrons. The fact is that this idea, so obvious in hindsight, escaped me.

There had been earlier opportunities, as discussed in Sections V and VI, to invent minority-carrier injection. As discussed for 4 April 1947, I had proposed a lightning arrestor composed of a sequence of very thin p-type and n-type layers in series. In a high electric field, the shallow potential hills would be flattened out so that carriers of either sign could go straight through: i.e., a form of electric breakdown would occur. During my work on diffusion theory of 24 April 1947, shown on Fig. 10, I had in hand all the necessary mathematical machinery needed to derive the p-n junction current–voltage formula that would have included injection by a forward-biased emitter junction but had simply done nothing about it. As shown in Fig. 11 for 16 September 1947, I had proposed a high-speed thermistor that also involved minority carriers passing over the potential energy maximum for them in a layer of the opposite conductivity type.

But for *none* of these ideas had I singled out the conception of injection of significant densities of minority carriers as potentially useful. By describing these lost opportunities in detail, I intend to illustrate how the path to creativity may be neither easy nor direct. My hope is that by being frank about my own limitations in seeing what later seemed so obvious, I may encourage my readers to persist rather than give up when they feel distressed over their own limitations.

23 January 1948 is the date for my completion of the conception of the junction transistor—the end of the path defined by the title of this historical note.

This completion of the concept resulted from an accident—but an accident for which my unsuccessful inventive efforts had prepared my mind. I shall first describe the conception as expressed in a notebook disclosure and then put it in perspective by reconstructing, as well as possible, the path that led to it.

Fig. 20 shows the first two pages of my five page disclosure of the junction transistor, written on a pad at home and later rubber-cemented into my notebook. The three key concepts discussed in Section II are all clearly expressed in Fig. 20: 1) exponentially increasing minority carrier injection across the emitter junction, 2) reverse bias on the collector junction, and 3) appropriate values both for dimensions and also for donor and acceptor densities. These structural considerations are discussed on the second page shown in Fig. 20 and are amplified on the third page (omitted from Fig. 20; for this, see [3, fig. 34], which starts with this paragraph:

> Some current will be drawn by control electrode. However, this will be small compared to the modulated current

so long as the concentration of holes in the p-layer is small compared to the concentration of electrons in the high concentration region to the left of the p-layer.

I followed the junction transistor disclosure with a disclosure of how to use transit time effects to obtain negative resistance—a line of thought stimulated by my indoctrination upon arriving at the Laboratories in 1936. The disclosures on these pages were witnessed four days later on 27 January by J. R. Haynes.

The delay of four days between disclosure and witnessing was longer than was typical for the transistor group during that period. The disclosure was written on Friday and witnessed on the following Tuesday. I am sure that if I had *really* appreciated the impact that the junction transistor would have, I would have driven the few miles necessary to obtain a witnessing signature the *same* day.

The title "High Power Large Area Semi-Conductor Valve" of the disclosure was *not* prophetic. The great contribution of the junction transistor has not been its high power handling capacity. Quite the contrary. The good ***high*** frequency performance at unprecedented *low* power levels is what revolutionized the computer industry. This revolution has included electronic switching in telephone exchanges, the objective that Mervin Kelly had emphasized to me with such enthusiasm during my first years or so at Bell Laboratories.

My lack of overwhelming enthusiasm at the time for the disclosure of Fig. 20 is understandable in terms of the path that led me to the concept of controlled minority carrier injection which was the essential addition that I made in the disclosure of Fig. 20. Two vivid memories and a notebook entry of mine of 20 February 1948 have enabled me to reconstruct the sequence of ideas about which the details have long since disappeared—and to do so with confidence in the accuracy.

My first vivid memory is that I was not undertaking a program intended to result directly in the invention of a transistor when the key concepts of Fig. 20 became apparent. Instead, I was trying to invent new experiments to improve the scientific aspects of the research on the p-type surface layer of the point-contact transistor.

The concept of imrefs (or quasi-Fermi levels as discussed in Section III) was inherent to these new experiments. Although I did not in 1948 use these words, both my vivid memory and my documentation support this reconstruction of this part of the history.

The supporting documentation is the description of the new experimental idea in my notebook entry proposal of 20 February 1948 that minority carriers near a forward biased p-n junction would have a different imref from the majority carriers surrounding them—and, most importantly, that this difference could be measured with suitable probes. These concepts are clearly—somewhat ungrammatically—expressed in that entry's review of what had led me to the disclosure of Fig. 20 in the following words:

> ... the potential should spread out as indicated. Near surface electron distribution may be in equilibrium with inside N [meaning underlying n-type material] & holes with point P. Hence probing surface with N-type point may give different value from P type point.

This was the type of thinking that had occupied me just before I wrote the disclosure of Fig. 20. The new experiments that I was trying to design would determine whether an n-type probe and a p-type probe would measure the same voltage at a given point on the surface. The following year, I included these considerations in my p-n junction paper [11] and introduced the term "internal contact potentials" as a parallel to differences in the potentials in space outside two different metals in contact with each other.

The origin of the junction transistor diagram of Fig. 20 probably arose in the course of planning imref experiments. As a preliminary to persuading my colleagues to conduct experiments, it was natural for me to "try simplest cases" in visualizing conceptual experiments. Simplest cases, like the "*L* Case" of Fig. 2, are often one-dimensional. A one-dimensional n-type probe would be an n-type layer making contact to the thin p-type layer. Such a mental picture would lead me to draw a diagram that would look very much like the junction transistor diagram of Fig. 20.

In effect, after doing this planning, the junction-transistor structure was staring me in the face, or at least, looking me squarely in my mind's eye. This mental picture stimulated the will to think about p-n junction amplifiers much as the 17 November 1947 breakthrough observation stimulated the will to think of Bardeen, Brattain, and Gibney about field-effect amplifiers. From that point, only a short step on the path of ideas would lead to the disclosure of Fig. 20.

A key new feature of the mental picture and its diagram of Fig. 20 was the accessibility of the middle layer to an external control voltage. This permitted its possible valve action in an amplifier.

In retrospect, it appears that I had wandered close to this idea on a number of occasions. For example, the n-type strip of 31 December 1947 did have its independent voltage source. But that voltage source did not act upon an interposed layer of one conductivity type lying between regions of opposite type. And in every case when I had thought of an interposed layer, I had not imagined making the layer accessible to external voltage control: The layer in the thermistor and those in the lightning arrester were all simply allowed to float at the voltages imposed on them by their neighbors.

My second vivid memory supports my belief in the accuracy of the foregoing reconstruction of my last steps on the path to the conception of the junction transistor. That second vivid memory is about a train ride in late January of 1948 from a meeting of the American Physical Society in New York City to my home in New Jersey. During that ride, I expounded to a fellow member of the transistor group—and did so, I feel, with great clarity, eloquence, and enthusiasm—about the possibility of creating new science about the point-contact transistor by measuring internal contact potentials near an emitter point—the imref concept of my ideas.

At the end of my discourse, I stopped and waited in anticipation for an enthusiastic endorsement. It didn't come. To my dismay, my colleague's honest evaluation was, in essence, nonsense. I was keenly disappointed. That made the memory vivid. This memory serves two purposes: 1) It confirms a date for the imref ideas; 2) It also shows that they weren't all that obvious after all.

That fact that my imref concepts in 1940 were hard to comprehend continued into at least 1954—and did so to my advantage. In 1954 I consulted with a company that wished to use a transistor switch to chop low dc voltages for measurements. I recalled my equations for floating emitter potentials in the "internal contact potential" section of the microwatt transistor paper [13]. On the basis of the imref results of that analysis, I proposed putting two n-p-n junction transistors back-to-back in a circuit that would open or close a path between the two emitter terminals. The collectors were connected together and so were the bases. Forward bias across the collector junctions closes the switch and reverse bias opens it. Imref theory was essential in predicting the good performance. The application, filed on 3 September 1954, issued as patent 2,891,171 on 16 June 1959. Imref concepts, although sometimes hard to teach, had engineering significance in 1954—and still do.

To return to 1948, I had a competitive urge to make some important transistor inventions on my own. I talked about "internal contact potentials" but kept the junction transistor concept pretty much under wraps for nearly a month. Then, in effect, my hand was forced by the observation of John Shive discussed in the next Section. Two days after that I wrote a notebook entry of 20 February 1948 to record the history. Twenty eight years later, I have used that entry as a key item in my reconstruction of the history presented in this Section VII.

VIII. From Conception to Realization of the Junction Transistor

The path from the conception to the realization of the junction transistor was no more direct than the path to its conception had been. I shall in this section focus on those aspects most closely related to the conception rather than to the final success of production technology which is discussed elsewhere, for example, in my 1973 article [3].

I have selected for comment six milestone developments on the path to the realization of the potential of the junction transistor that confirmed, or extended, or realized, basic junction transistor concepts. In chronological order, insofar as possible, these are:

1) The extension of the analysis and concepts of possible devices. This applies to the filing of the junction transistor patent application [14] on 26 June 1948. At that time I tried to put anything potentially useful, but not so wild as to be ridiculous, into the specification. I included heavy doping near contacts, heterojunctions with wide energy gaps to increase emitter efficiency, transit time effects for negative resistance (an idea motivated by my early exposure to the vacuum tube department), many layer structures for modulation, etc. My p-n junction article [11], published in 1949, introduced imrefs ("nonequilibrium quasi Fermi levels" in [11, eq. (2.4)] and analyzed minority carrier injection in detail. The theoretical possibilities presented in these analyses did stimulate the will to think about how to realize the desired structures.

2) Chiefly during the second half of 1948, the reality of minority carrier injection was established. I shall discuss this key topic in detail after completing the brief discussion of 3) through 6).

3) In April of 1949, an "existence proof" germanium junction transistor gave power gain using both emitter and collector p-n junctions (see the *7 April 1949* item in Table I and especially [3] for details).

4) In 1950, a good p-n junction was grown by G. K. Teal, M. Sparks, and E. Buehler by changing the doping of a melt as a crystal was grown. This was an outgrowth of the 1918 Czochralski method of pulling crystals from the melt, the 1950 Teal–Little application to germanium, J. A. Morton's decision to support Teal's enthusiasm for growing single crystal germanium, and Teal's addition of impurities to the melt. (For details see [3] and [6] and Teal's patent 2,703,296 issued 1 Mar 55, filed 15 Jun 50; Teal's laboratory notebooks for the period of this historical note were missing from the Laboratories' files.) The characteristics of this junction were in good quantitative accord with the theory of [11] to a degree previously unprecedented for a semiconductor rectifier. This was a significant confirmation of the theory for junction transistors. (For references, see F. S. Goucher *et al.* in [6], [13] and the article by Teal in this bicentennial issue.)

5) In *April of 1950*, "double doping" of a melt was used to grow an n-p-n structure to make a grown-junction transistor. One was demonstrated on 20 April 1950 according to my marginal note in Fig. 20. This nonphotogenic device did perform according to theory but had a wide base, and poor frequency response and provoked little interest. (See Figs. 21, 22, and 23 and [3]). For about nine months, the efforts to improve junction transistors were practically negligible at the Laboratories.

6) The will to think about exploiting the junction transistor's potential was aroused in *early 1951* by two factors: my interest in promoting the use of transistors in proximity fuzes and R. L. Wallace's recognition of their potential advantages for such purposes if the base layer could be made sufficiently thin. The resultant achievement of microwatt transistors then provoked the will to think of practical junction transistor development. (See Fig. 24 and also [3], the Shockley, Sparks, and Teal article [13], and in [6], W. J. Pietenpol and R. L. Wallace.)

I regard 2), the demonstration that minority carrier injection could occur and be really useful, to be a key development in 1948. Actually, injection of minority carriers by rectifying, metal point-contacts had been observed at least once, and possibly twice, before my disclosure of injection by p-n junctions in Fig. 20. In neither case, had it been interpreted correctly.

The earlier of these two cases was reported in May 1947 by a group at Purdue University. Ralph Bray and his colleagues had proposed that high electric fields lowered the

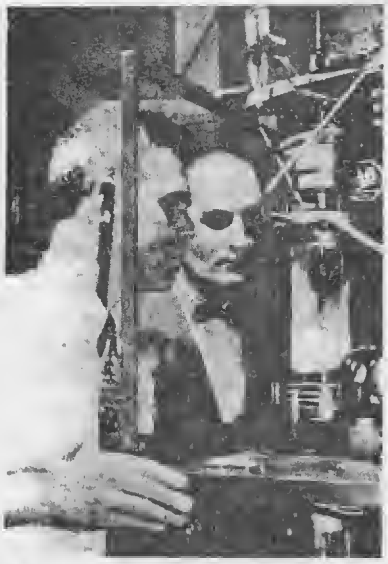
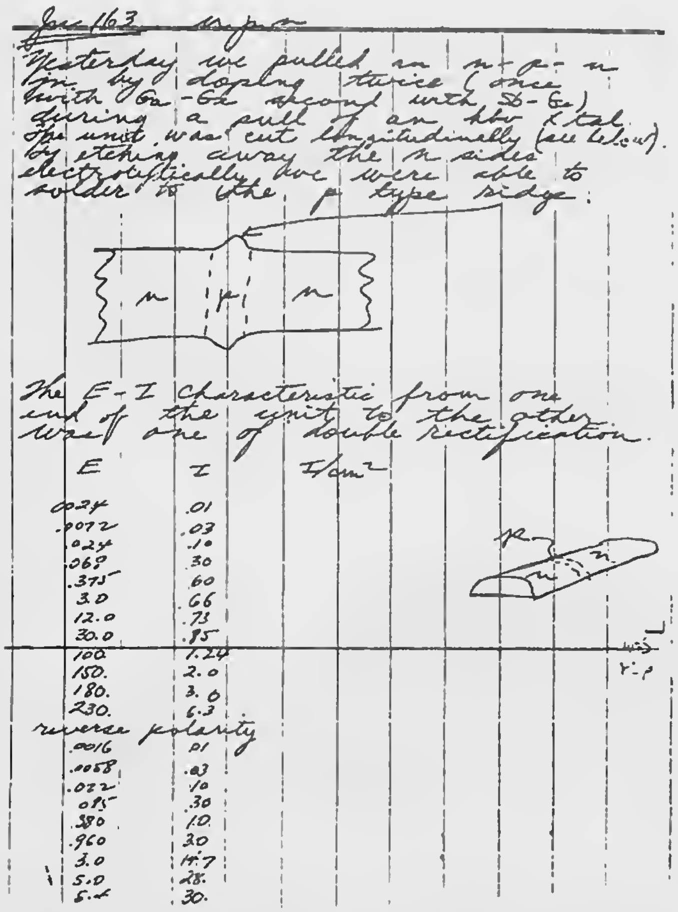


Fig. 21. A crystal-growing apparatus of the form developed by G. K. Teal being observed by E. Buehler, who grew most of the crystals used in the semiconductor development and research programs at Bell Laboratories for many years, and M. Sparks who fabricated the first of the good junction transistors.



Yesterday we pulled an n-p-n jn. by doping twice (once with Ga-Ge second with Sb-Ge) during a pull of an hbv xtal. The unit was cut longitudinally (see below). By etching away the n sides electrolytically we were able to solder to the p type ridge.

The E-I characteristic from one end of the unit to the other was one of double rectification.

E	I
.0024	.01
.0072	.03
.024	.10
.068	.30
.375	.60
3.0	.66
12.0	.73
30.0	.85
100.	1.24
150.	2.0
180.	3.0
230.	6.3
reverse polarity	
.0016	.01
.0058	.03
.022	.10
.085	.30
.380	1.0
.960	3.0
3.0	14.7
5.0	28.
5.4	30.

Fig. 22. A disclosure of 12 April 1950 by Morgan Sparks of his technique for making electrical connection to a base layer in fabricating a junction transistor.

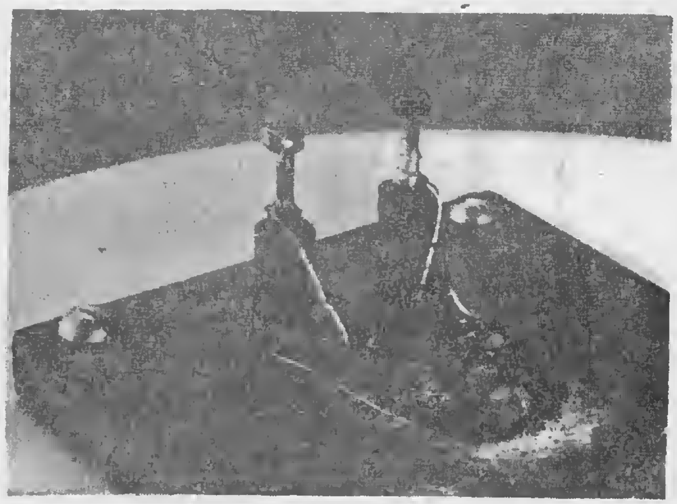


Fig. 23. The first successful sandwich-structure junction transistor.

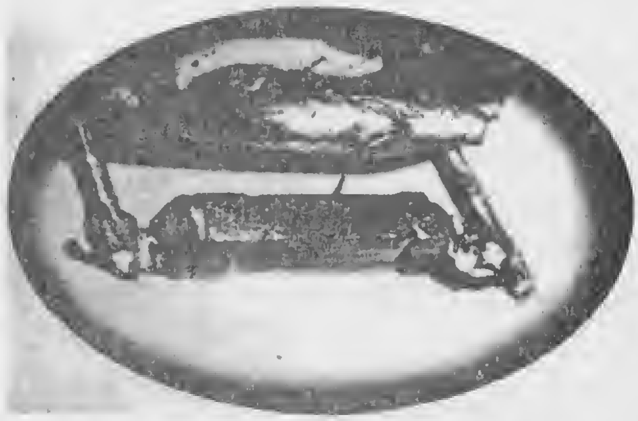


Fig. 24. An example of the first microwatt junction transistor, the device that may be said to have launched the transistor era.

resistivity of high back-voltage germanium (see [4] for details and references). What really increased the conductivity was the injection of minority holes, plus neutralizing electrons.

The second case may have occurred when the point-contact transistor was born. About this event, in "A Reflection on the Discovery 27 Years Later" [15], Brattain wrote:

> . . . Bardeen and I were simply trying to make a good "Field Effect" device and as a result we were put in a position to observe, for the first time, a phenomenon (namely minority carrier injection) now called the "Transistor Effect"—and to use this to make a transistor!

Brattain's description does not seem to me to be adequately qualified as to the facts at the time. When the "transistor effect" was observed in the sense of the birth of the point-contact transistor, no one had any idea that minority-carrier injection was being observed any more than had been true for Bray's results. So far as I know, the first suggestion that minority carrier injection might be an important mechanism was made five weeks later in my disclosure of Fig. 20. Indeed, so far as the "Transistor Effect" in the point-contact transistor is concerned, at the time of the public announcement of the transistor more than six months later, the explanation of the "transistor effect" emphasized the p-type surface layer. For example,

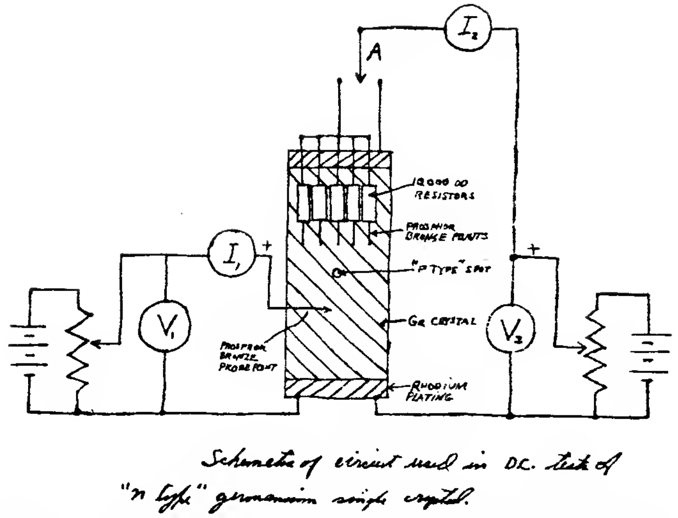


Fig. 25. J. R. Haynes' notebook entry of 6 June 1948 showing the first version of the Haynes-Shockley injected-carrier drift-velocity experiment.

an article based on the announcement and coauthored by Donald G. Fink, then editor-in-chief of *Electronics,* described the emitter point thus [20]:

> In a Transistor, the positive point contact causes the release of holes in the surface layer of the germanium, which is prepared in a similar manner to a high back-voltage rectifier. These holes spread away from the point, flowing in all directions along the surface (but not into the body of the semiconductor) . . . .

This description is consistent with the somewhat less specific earliest publications by Bardeen and Brattain [8], [9] and is also in accord with their 1949 transistor article ([4] page 1211): "The early experiments suggested flow along the surface." I am not now certain whether a surface layer or injection into the bulk was what made the first point-contact transistor amplify. So far as I have found, the word "injection" first appeared in journals in my publications with E. J. Ryder and J. R. Haynes as discussed below.

Striking experimental evidence for injection by an emitter point was reported on *18 February 1948* by John Shive at one of the irregularly held confidential conferences restricted to the transistor group at Bell Laboratories. Shive had brought emitter and collector points very close together in a new way—on opposite sides of a very "thin wedge or slab"—and had obtained transistor action. Shive was not then aware of my **23 January 1948** disclosure and his interpretation did not include the concept of minority carriers diffusing in the presence of majority carriers which were essentially in equilibrium. At the conference, as soon as I had heard Shive's report, I presented the ideas of my junction transistor disclosure and used them to interpret Shive's observation. I discussed these events more fully in my 1973 paper [3].

A classic semiconductor laboratory experiment resulted from my determination to solve the puzzle of whether the point-contact transistor did operate primarily by injection or by surface layer conduction. The outcome of this was the decisive Haynes–Shockley experiment. The earliest relevant notebook entry that I found appears in Fig. 25. This dc experiment was later modified to measure the transit time of injected holes to a movable detector probe consisting of a collector point. (These measurements were the second time that J. R. Haynes and I had collaborated in measuring a drift velocity, the first being that of electrons in silver chloride, see my book [12].) We verified that the holes moved through the bulk by putting the emitters and collector on opposite sides. These experiments also led to measurements of surface recombination velocities.

Still another experiment established that substantial reduction of bulk conductivity could be produced by hole injection. This experiment was carried out by E. J. Ryder.

More detailed descriptions of these experiments and references will be found in my Nobel Lecture [6], my 1973 article [3], and in Bardeen and Brattain's point-contact transistor article [4].

IX. Conclusion

From the evidence reported above gathered by my operations research on the history related to the conception of the junction transistor, the reader may draw many conclusions. One of them is that the motivations were of a mixed variety—both practical and scientific. Rather than enumerate a definite set of conclusions, I shall discuss some philosphical issues that are broader and underlie the specific topics covered in this historical note. For this purpose, I shall propose and give answers to several frequently arising questions.

Was the transistor a product of an engineering program focused on a practical goal? Or was the transistor a by-product of pure research unsullied by any motivation other than a search for knowledge?

I liken these questions to the old bromide: "Have you ceased beating your wife? Answer 'Yes' or 'No.' "

What actually went on was a mixture which fits into the pattern of "creative-failure methodology" which was what we intuitively put into action when frustrated by the failure of the field-effect experiments. Bardeen's concept of surface states as shielding the interior from external fields gave a practical significance to what had before been largely a theoretical concept. Both Igor Tamm in Russia (co-winner of the 1958 Nobel prize in physics for work on Cerenkov radiation) and I had done theoretical calculations showing that surface states should exist. However, no significant observable implications of these had been proposed. Indeed our surface states were of a highly mathematical nature and related to perfect crystalline surfaces. In contrast, Bardeen's concept of surface states was more empirical. His states were undefined in terms of their quantum mechanical origin and allowed for the possibility that defects on the surface might contribute to their existence—concepts in harmony with knowledge about the states due to donors and acceptors in the interior of the semiconductor.

Bardeen quickly recognized that his surface states had broad implications and were not restricted solely to explaining the field-effect failure. His surface states also resolved a number of mysteries about semiconductor surfaces including their rectifying characteristics when

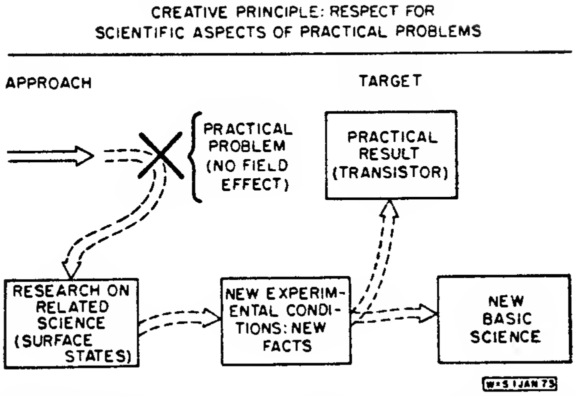


Fig. 26. Respect for the scientific aspects of practical problems—a feature of the creative-failure methodology that led to the point-contact transistor.

contacted either by metal points or else by other semiconductors—a dramatic example of creative-failure methodology in action.

Our semiconductor research team abandoned efforts to make a field-effect transistor and Brattain and Bardeen in particular emphasized research on new science related to Bardeen's surface states. We intuitively applied a feature of creative-failure methodology that some years later I analyzed and defined as *respect for the scientific aspects of practical problems.*

Fig. 26 represents how this principle operates in general. The phrases in parentheses interpret its vital role in creating the transistor: The attempt to make a semiconductor amplifier was blocked by a practical problem. But how we responded to this failure to reach the practical result was creative. The failure stimulated Bardeen's creative suggestion of the surface states. Our research on the science related to the surface states led to new experimental conditions and to new observations. Then the breakthrough observation of 17 November 1947 showed at long last how to overcome the blocking effect of the surface states. This new possibility motivated "the will to think" phenomena that led to the peak of creativity that followed immediately thereafter.

There may be today, 28 years after the invention of the transistor, an ironic aspect of the emphasis that I have given to "respect for the scientific aspects of practical problems" as an important creative principle in industrial research. By assigning so much emphasis to this feature of creative-failure methodology—a feature that I helped to establish—I may have become out-dated and be reflecting attitudes that are more appropriate to the experiences that I recall of the 1940's than they are to industrial research today.

In 1946 when the semiconductor research group focused on the basic science, leaders of some other research department groups urged me to emphasize *practical* semiconductor difficulties in the telephone plant. Our group was of one mind and we followed the wise course of working, not upon such practical but messy semiconductors as selenium, copper oxide, and nickel oxide, but instead on the best understood semiconductors of all—silicon and germanium—a "try simplest cases" approach.

For these best understood semiconductors, not [illegible] the theoretical concepts developed, as discussed in Se[illegible]n V, largely during World War II, had been experimentally verified. We elected to concentrate on the remaining gaps, among these being the recently proposed surface states. We felt that it was better to understand these two simplest elemental semiconductors in depth rather than to attempt to add piecemeal contributions to a variety of other materials.

In assigning our highest priority to the primarily scientific aspects, we chose those related to the problems that blocked our approach to the long-range practical goal—the creation of a semiconductor amplifier, later to be called the "transistor."

My feeling of being out of date occurred after I had stressed Fig. 26 at Bell Laboratories during a rehearsal in February 1973 of a lecture version of [3]. During a subsequent luncheon conference, a young scientist told me that he was puzzled that I had so emphasized "respect for the scientific aspects of practical problems." The approach that I endorsed so vehemently seemed so natural to him that it scarcely called for any emphasis at all. I felt separated from his appraisal by a "generation gap": What in 1947 had been, in the eyes of at least some colleagues, a pioneering advance from the Edisonian methods of trial and error to achieve practical goals was now, 25 years later, taken for granted. And the transistor story had probably helped to bring this about. By giving such a strong sales pitch to what today no longer needed selling, I had made my words become an echo of the past.

The additional question that I shall use to introduce some general observations related to the themes of this historical note was asked by a senator at the hearing on economic growth mentioned in footnote 1 of Section I: "Is there something unique about Bell Laboratories as compared to other very large corporations?"

I shall give two answers, the first being a reference to a remarkable article written by Richard R. Nelson, then a member of the RAND Corporation [21]. Nelson carried out operations research on Bell Laboratories records and interviewed a number of the personnel involved. I was not there at the time and do not remember any contact with Nelson. However, his grasp of the history and of the role of managerial policy is the best that I have seen published. I recommend it highly.

My second answer is a revised paraphrase of part of my response in the testimony of footnote 1 in Section I:

I believe that there are significant comments to make about motivational inducements, or their absence, at Bell Laboratories. The payments for patents at Bell Laboratories was, and I think still is, $1.00 paid in advance for signing the patent agreement while becoming employed. In contrast to that, if one tried to set up a set of rules whereby one would be able to determine the just deserts for contributions to invention and innovation by the process of examining what existed on records on paper, then I believe that the effects in a place like Bell Laboratories would be very adverse. It would leave out the most important feature of sound human judgment by competent

administrators. (For more details on my reasoning see [22].)

Cumulative *wisdom*, based on competence and experience, is going to be better than any *rule* that can be set up and put on a computer in an attempt to give to each person his just rewards for his contributions to creativity. The organization and the progression of people through Bell Laboratories and the tenure of individuals there—these are all of high caliber.

A few other places must be comparable in quality of personnel and continuity of experience, but I believe that Bell Laboratories must be very near the top. This continuity in personnel and spirit means that one can count on about the best that can be expected from management considering the fundamental human limitations that exist everywhere. Thus, in a good organization of capable experienced people, an employee can count upon a really high degree of justice. For example, I think of a person who might promote high inventiveness among the individuals in the group he led so that they would file many valuable pattents while he, their leader, did not. Yet he might be the one most responsible for the group creativity. In a good organization, administrative judgment would give appropriate recognition and this would be felt, although perhaps not analyzed, by all who were involved.

The form of managerial justice that I am trying to describe is more difficult to codify so as to approach administration by mechanized rules than it is to carry out in practice by proven competent leaders. About the excellence of Bell Labs, I have often speculated about the extent of the influence of the humanitarian motives of Alexander Graham Bell. Bell attempted to better the lot of humanity generally by publishing his genetic research on longevity, deafness, and how to improve the race. Bell's picture as honorary president of the Second International Congress of Eugenics in 1921 is the frontispiece of its Proceedings. To what degree, I wonder, do these basic humanitarian motivations and values of Alexander Graham Bell continue to permeate the system and to contribute to its outstanding contributions to the benefit of people and to economic growth?

I shall close with a direct quotation from the testimony of footnote 1 of Section I, that I have paraphrased in part in the above paragraphs. In it I tried to express my interpretation of the spirit of creative-failure methodology:

"What I say about myself—and I am sure most creative people would say the same thing—is that, when we look at how long it took us to get certain ideas, we are impressed with how dumb we were—on how long it took us, and how stupid we were. But we have learned to live with this stupidity, and to find from it what relationships we should have seen in the first place. This recognition that we aren't perfect but that persistence pays is a very important factor, I think, in giving one *the will to think*—you don't need to worry so much about the mistakes you make, providing they are not too dangerous or too expensive."

References

[1] W. Shockley, "Transistor history, applied research and science teaching," invited lecture, Japanese Institution of Electrical Engineers, 75th Anniversary Meeting, Tokyo, Japan, Oct. 1963, published in *J. IEE Japan*, 84–2, No. 905, 147.

[2] ——, "The invention of the transistor: an example of creative-failure methodology," in *European Solid State Device Research Conference, 2nd* (University of Lancaster), 1972, (Paper 4), pp. 55–75. *Solid-State Devices*, 1972; *Proceedings of the Second European Solid State Device Research Conference*, (London, Institute of Physics), 1973, Conference Series #15.

[3] ——, "The invention of the transistor—An example of creative-failure methodology,' " National Bureau of Standards Special Publication 388, Proc. of Conf. on the Public Need and the Role of the Inventor, Monterey, Calif., pp. 47–89, Jun. 11–14, 1973 (issued May 1974).

[4] J. Bardeen and W. Brattain, "Physical principles involved in transistor action," *Phys. Rev.*, vol. 75, pp. 1208–1225, 1949.

[5] J. Bardeen, "Semiconductor research leading to the point contact transistor," Nobel Lecture, Dec. 11, 1956, published in *Les Prix Nobel*, pp. 77–99, 1957.

[6] W. Shockley, "Transistor technology evokes new physics," Nobel Lecture, Dec. 11, 1956, published in *Les Prix Nobel*, pp. 100–129, 1957.

[7] W. Brattain, "Surface properties of semiconductors," Nobel Lecture, Dec. 11, 1956, published in *Les Prix Nobel*, pp. 130–137.

[8] J. Bardeen and W. H. Brattain, "The transistor—A semiconductor triode," *Phys. Rev.*, vol. 74, p. 230, 1948.

[9] W. H. Brattain and J. Bardeen, "Nature of the forward current in germanium point contacts," *Phys. Rev.*, vol. 74, p. 231, 1948.

[10] W. Shockley and G. L. Pearson, "Modulation of conductance of thin films of semi-conductors by surface charges," *Phys. Rev.*, vol. 74, p. 232, 1948.

[11] W. Shockley, "The theory of p-n junctions in semiconductors and p-n junction transistors," *Bell Syst. Tech. J.*, vol. 28, pp. 435–489, 1949.

[12] ——, *Electrons and Holes in Semiconductors*. New York: Van Nostrand, 1950.

[13] ——, M. Sparks, and G. K. Teal, "p-n junction transistors," *Phys. Rev.*, vol. 83, pp. 151–162, 1951.

[14] ——, U.S. Patent 2,569,347, issued Sept. 25, 1951 (filed June 26, 1948).

[15] W. H. Brattain, "Discovery of the transistor effect," *Adventures in Experimental Physics*, ϵ, vol. 5, pp. 1–31, 1976.

[16] W. Shockley, "Articulated science teaching and balanced emphasis," *IEEE Spectrum*, vol. 3, pp. 49–58, June. 1966.

[17] ——, "THINKING about THINKING improves THINKING," *IEEE Student Journal*, pp. 11–16, Sept. 1968.

[18] ——, "A 'try simplest cases' resolution of the Abraham-Minkowski controversy on electromagnetic momentum in matter," in *Proc. Nat. Academy of Sciences*, vol. 60, no. 3, pp. 807–813, July 1968.

[19] W. H. Brattain, (statement on 25th Anniversary of Transistor), *Bell Laboratories Record*, vol. 50, p. 338, 1972.

[20] D. G. Fink and F. H. Rockett, "The transistor—a crystal triode," *Electronics*, Sept. 1948.

[21] R. R. Nelson, "The link between science and invention: the case of the transistor," in *The Rate and Direction of Inventive Activity: Economic and Social Factors*. Princeton, NJ: Princeton University Press, pp. 549–583, 1962.

[22] W. Shockley, "On the statistics of individual variations in productivity in research laboratories," *Proc. IRE*, vol. 45, no. 10, p. 1409, 1957.